

Wright State University

CORE Scholar

---

Biological Sciences Faculty Publications

Biological Sciences

---

5-23-2007

## Essential Elements of a Defense-Review of DNA Testing Results

Dan E. Krane

Wright State University - Main Campus, dan.krane@wright.edu

Follow this and additional works at: <https://corescholar.libraries.wright.edu/biology>



Part of the [Biology Commons](#), [Medical Sciences Commons](#), and the [Systems Biology Commons](#)

---

### Repository Citation

Krane, D. E. (2007). Essential Elements of a Defense-Review of DNA Testing Results. .  
<https://corescholar.libraries.wright.edu/biology/254>

This Presentation is brought to you for free and open access by the Biological Sciences at CORE Scholar. It has been accepted for inclusion in Biological Sciences Faculty Publications by an authorized administrator of CORE Scholar. For more information, please contact [library-corescholar@wright.edu](mailto:library-corescholar@wright.edu).

# Essential elements of a defense-review of DNA testing results

Dan E. Krane, Wright State University, Dayton, OH

Forensic Bioinformatics  
([www.bioforensics.com](http://www.bioforensics.com))

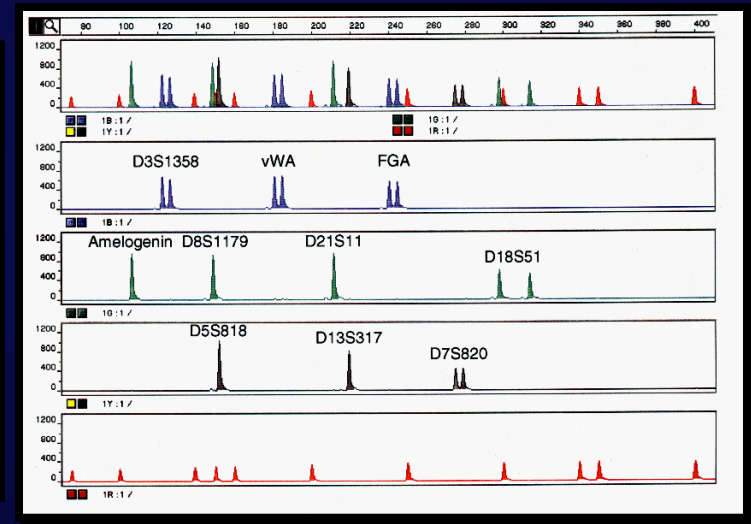
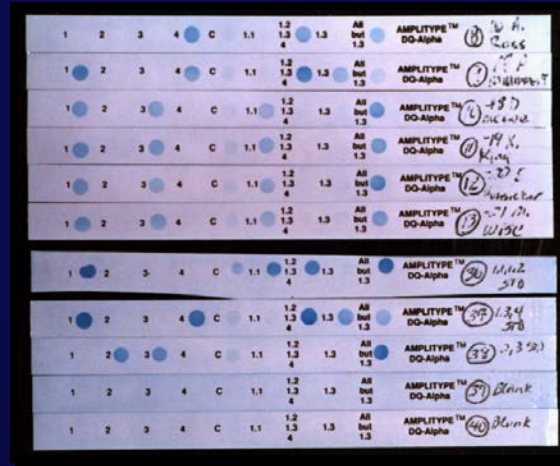
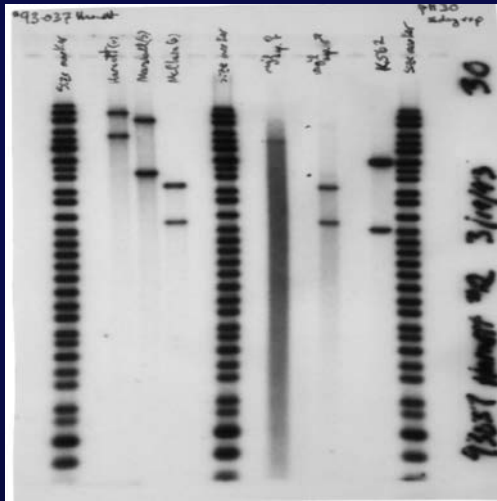
The science of DNA profiling is  
sound.

But, not all of DNA profiling is  
science.

**RFLP**  
**AUTORAD**  
**Allele = BAND**

**DQ-alpha**  
**TEST STRIP**  
**Allele = BLUE DOT**

# Automated STR ELECTROPHEROGRAM Allele = PEAK

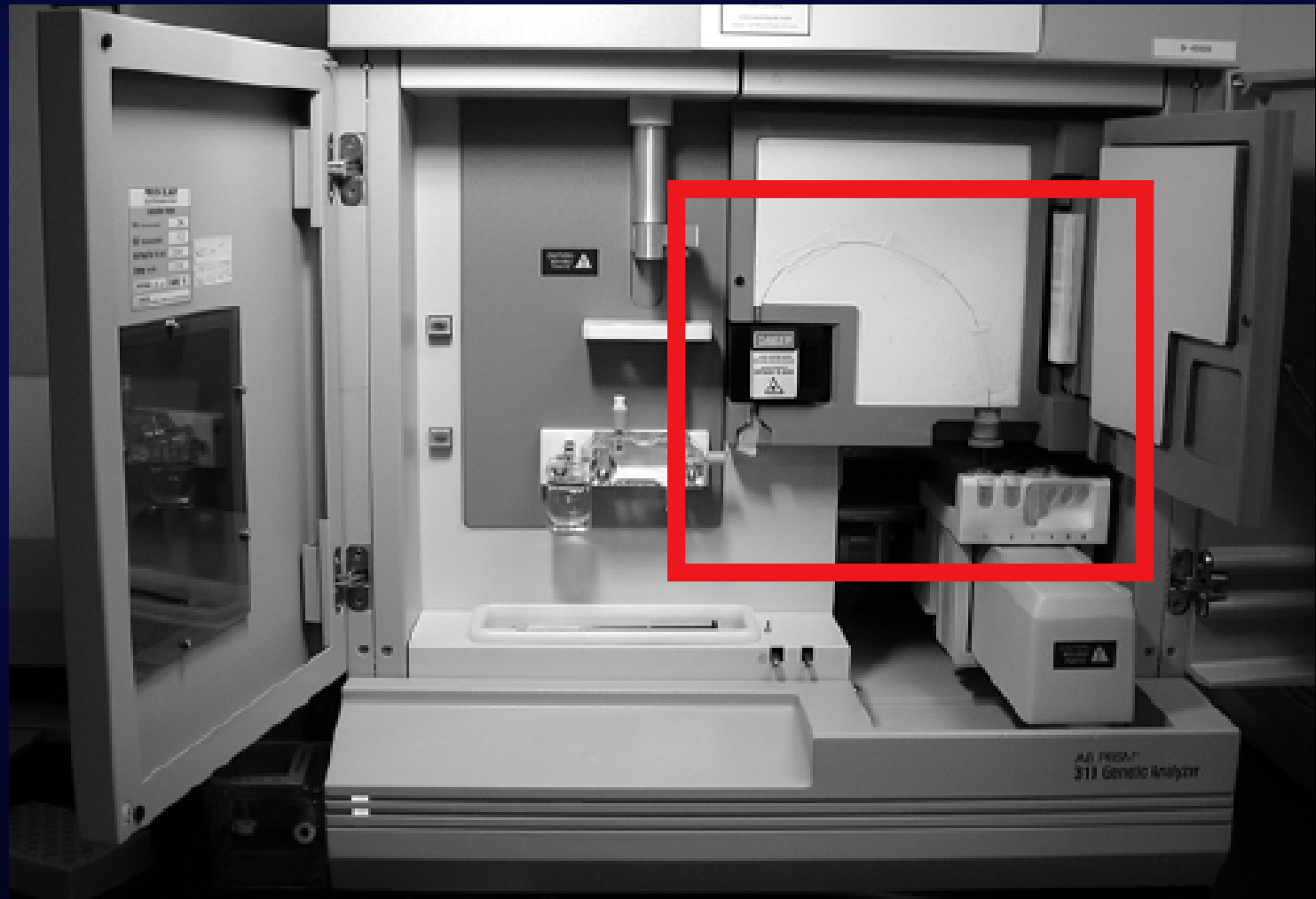


# DNA content of biological samples:

| Type of sample                  | Amount of DNA   |
|---------------------------------|-----------------|
| Blood                           | 30,000 ng/mL    |
| stain 1 cm <sup>2</sup> in area | 200 ng          |
| stain 1 mm <sup>2</sup> in area | 2 ng            |
| Semen                           | 250,000 ng/mL   |
| Postcoital vaginal swab         | 0 - 3,000 ng    |
| Hair                            |                 |
| plucked                         | 1 - 750 ng/hair |
| shed                            | 1 - 12 ng/hair  |
| Saliva                          | 5,000 ng/mL     |
| Urine                           | 1 - 20 ng/mL    |

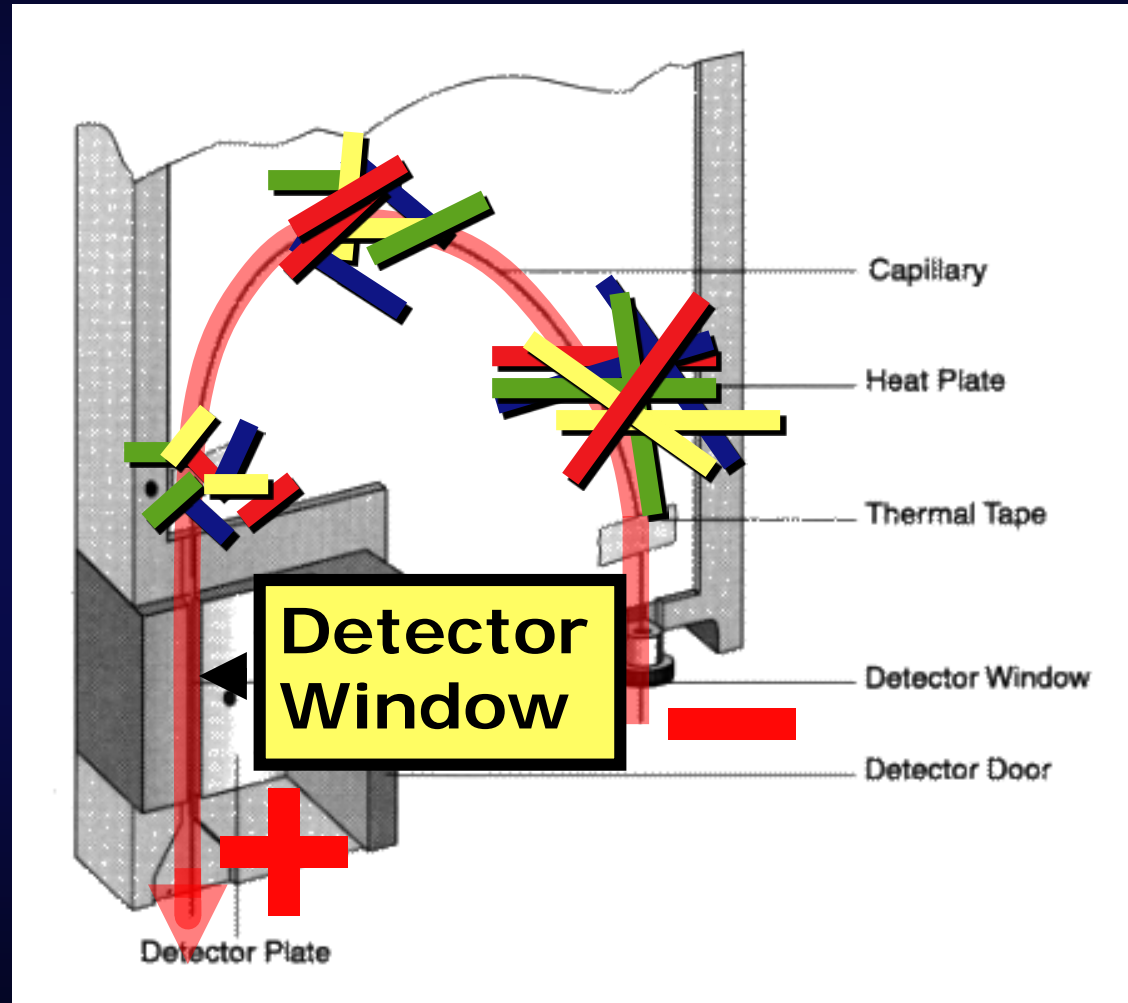
# Automated STR Test

# The ABI 310 Genetic Analyzer



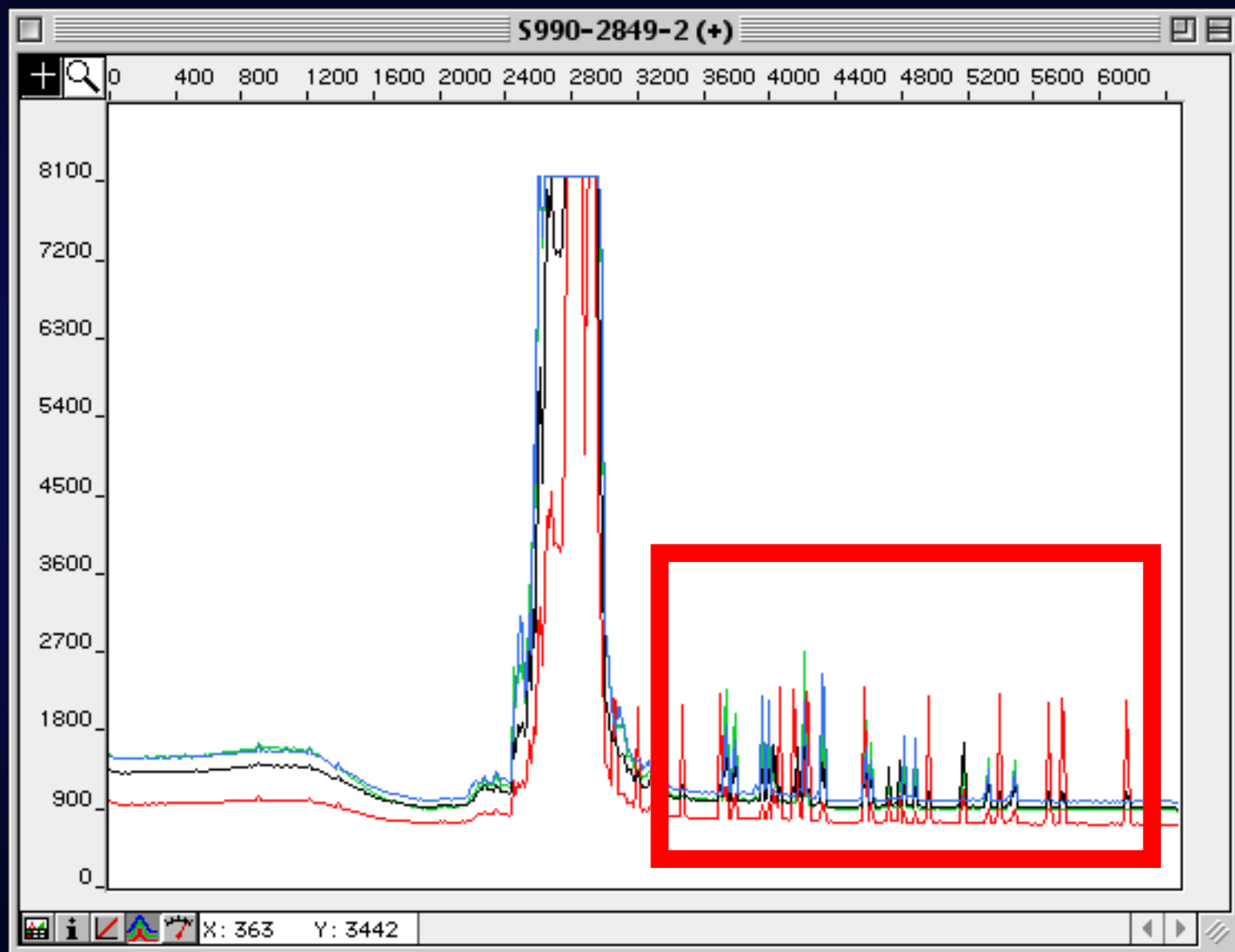
# ABI 310 Genetic Analyzer: Capillary Electrophoresis

- Amplified STR DNA injected onto column
- Electric current applied
- DNA pulled towards the positive electrode
- DNA separated out by size:
  - Large STRs travel slower
  - Small STRs travel faster
- Color of STR detected and recorded as it passes the detector





# Profiler Plus: Raw data



# Statistical estimates: the product rule

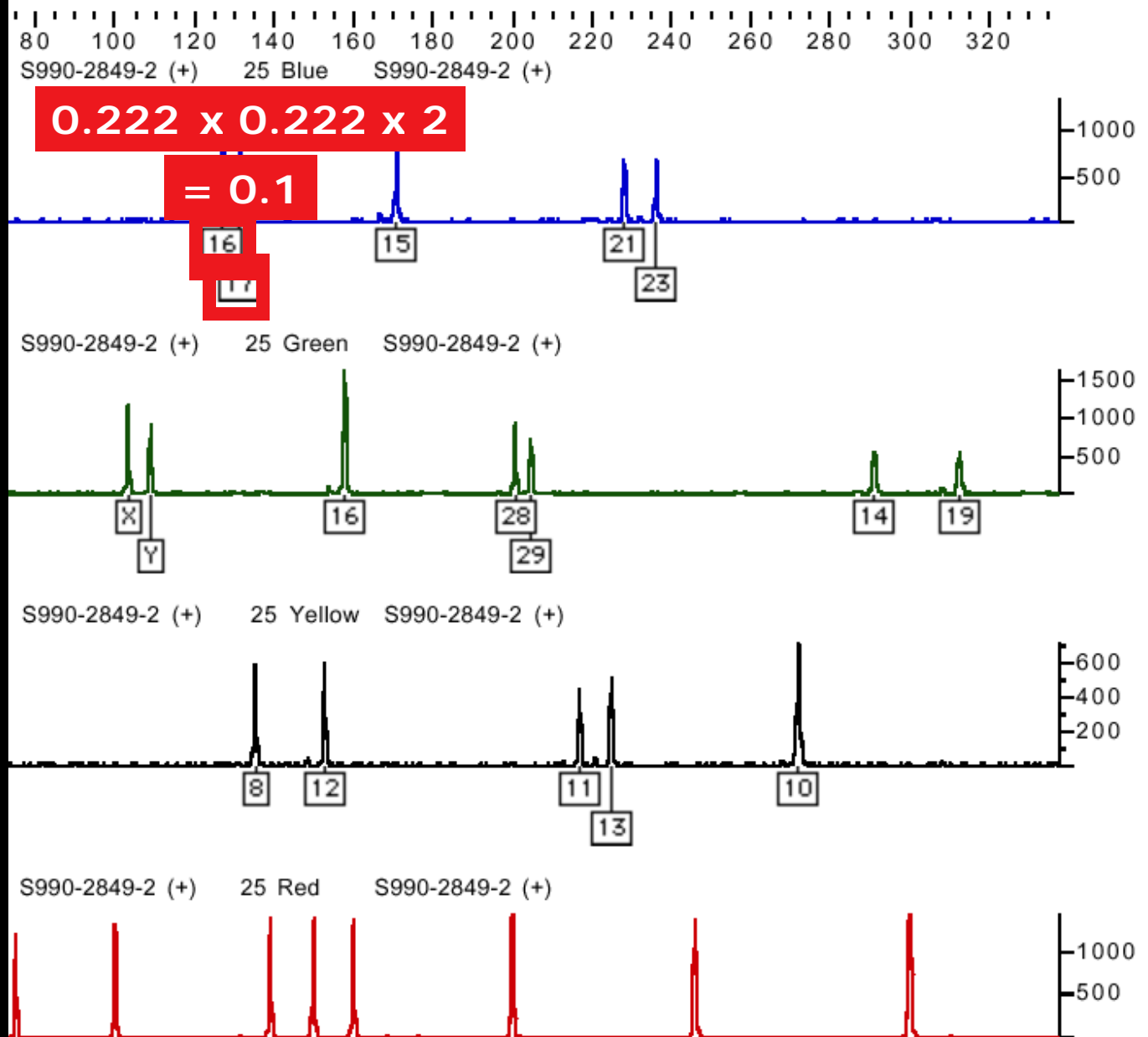
## Allele Frequencies

Locus D3S1358  
Race Caucasian  
(N = 203)

| Allele | Frequency |
|--------|-----------|
| 12     | 0.012     |
| 13     | 0.012     |
| 14     | 0.140     |
| 15     | 0.222     |
| 16     | 0.222     |
| 17     | 0.222     |
| 18     | 0.103     |
| 19     | 0.012     |

Locus vWA  
Race Caucasian  
(N = 196)

| Allele | Frequency |
|--------|-----------|
| 11     | 0.012     |
| 12     | 0.012     |
| 13     | 0.012     |
| 14     | 0.102     |
| 15     | 0.082     |



# Statistical estimates: the product rule

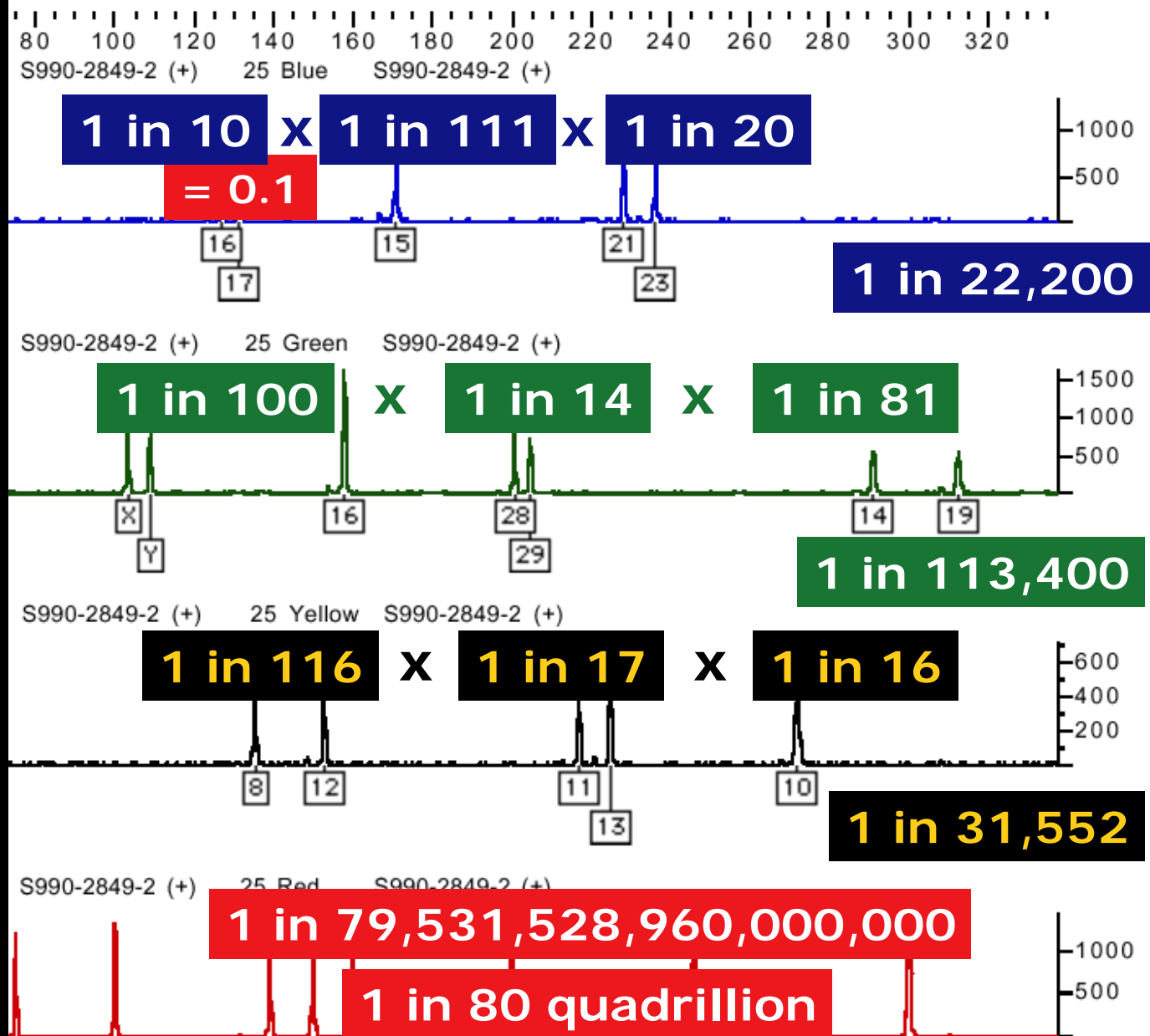
## Allele Frequencies

Locus D3S1358  
Race Caucasian  
(N = 203)

| Allele | Frequency |
|--------|-----------|
| 12     | 0.012     |
| 13     | 0.012     |
| 14     | 0.140     |
| 15     | 0.246     |
| 16     | 0.222     |
| 17     | 0.222     |
| 18     | 0.163     |
| 19     | 0.012     |

Locus vWA  
Race Caucasian  
(N = 196)

| Allele | Frequency |
|--------|-----------|
| 11     | 0.012     |
| 12     | 0.012     |
| 13     | 0.012     |
| 14     | 0.102     |
| 15     | 0.082     |



What more is there to say after you  
have said: "The chance of a  
coincidental match is one in 80  
quadrillion?"

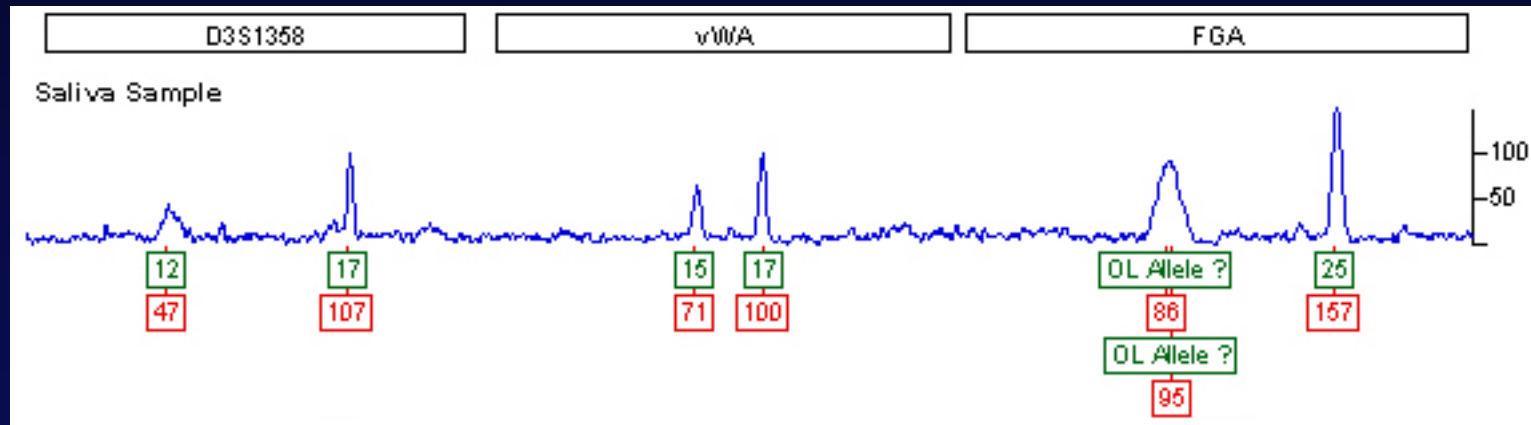
What more is there to say after you have said: "The chance of a coincidental match is one in 80 quadrillion?"

- Two samples really do have the same source
- Samples match coincidentally
- An error has occurred

The science of DNA profiling is  
sound.

But, not all of DNA profiling is  
science.

# Opportunities for subjective interpretation?



Can "Tom" be excluded?

Suspect

Tom

D3

17, 17

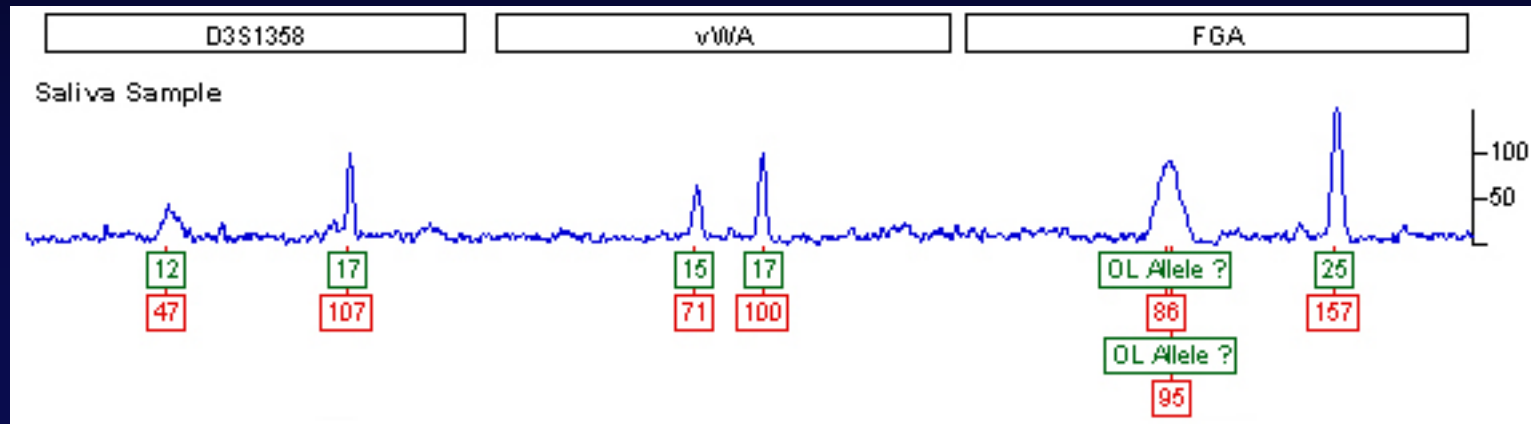
vWA

15, 17

FGA

25, 25

# Opportunities for subjective interpretation?



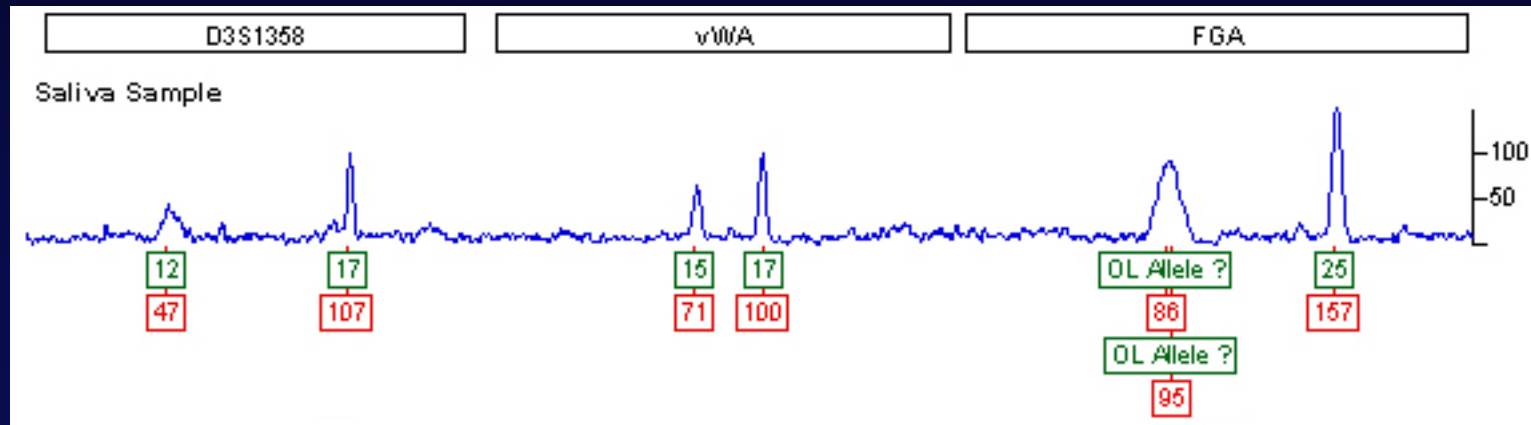
Can "Tom" be excluded?

| <u>Suspect</u> | <u>D3</u> | <u>vWA</u> | <u>FGA</u> |
|----------------|-----------|------------|------------|
| Tom            | 17, 17    | 15, 17     | 25, 25     |

No -- the additional alleles at D3 and FGA are "technical artifacts."



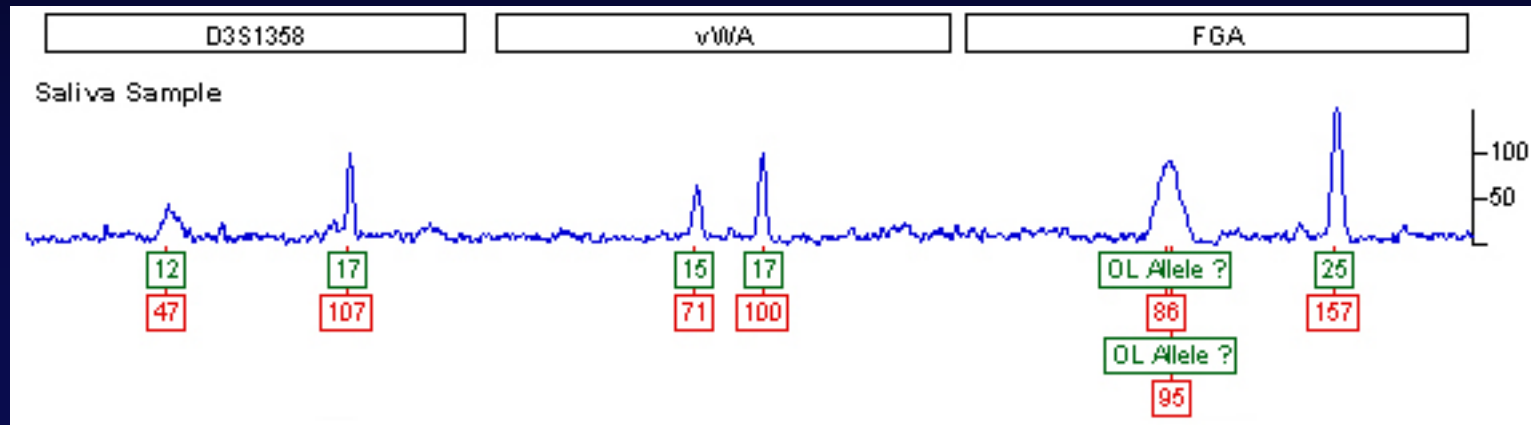
# Opportunities for subjective interpretation?



Can "Dick" be excluded?

| <u>Suspect</u> | <u>D3</u> | <u>vWA</u> | <u>FGA</u> |
|----------------|-----------|------------|------------|
| Tom            | 17, 17    | 15, 17     | 25, 25     |
| Dick           | 12, 17    | 15, 17     | 20, 25     |

# Opportunities for subjective interpretation?

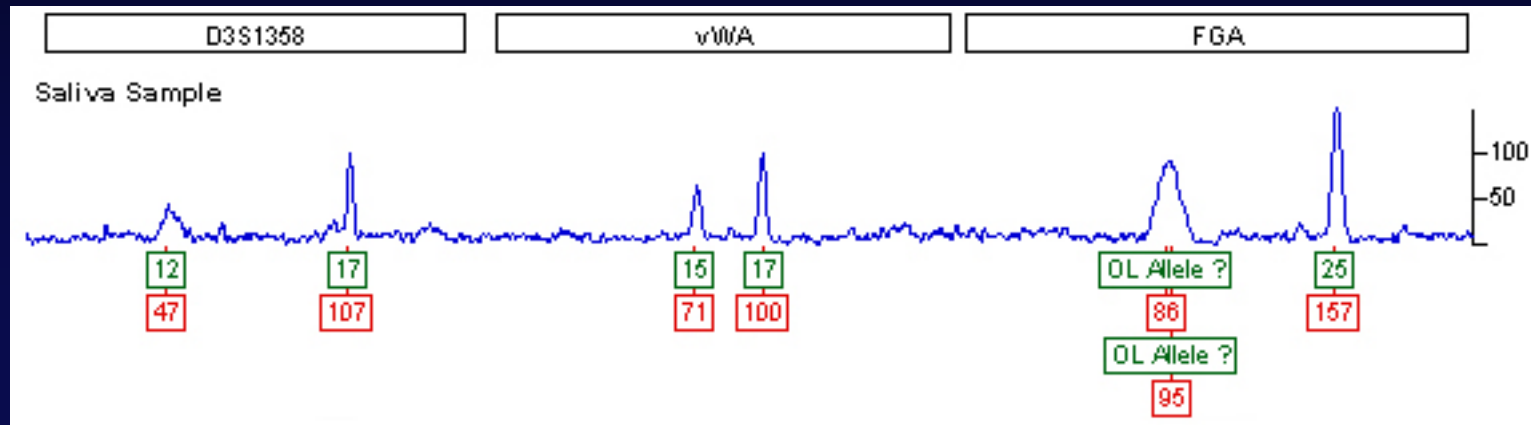


Can "Dick" be excluded?

| <u>Suspect</u> | <u>D3</u> | <u>vWA</u> | <u>FGA</u> |
|----------------|-----------|------------|------------|
| Tom            | 17, 17    | 15, 17     | 25, 25     |
| Dick           | 12, 17    | 15, 17     | 20, 25     |

No -- stochastic effects explain peak height disparity in D3; blob in FGA masks 20 allele.

# Opportunities for subjective interpretation?

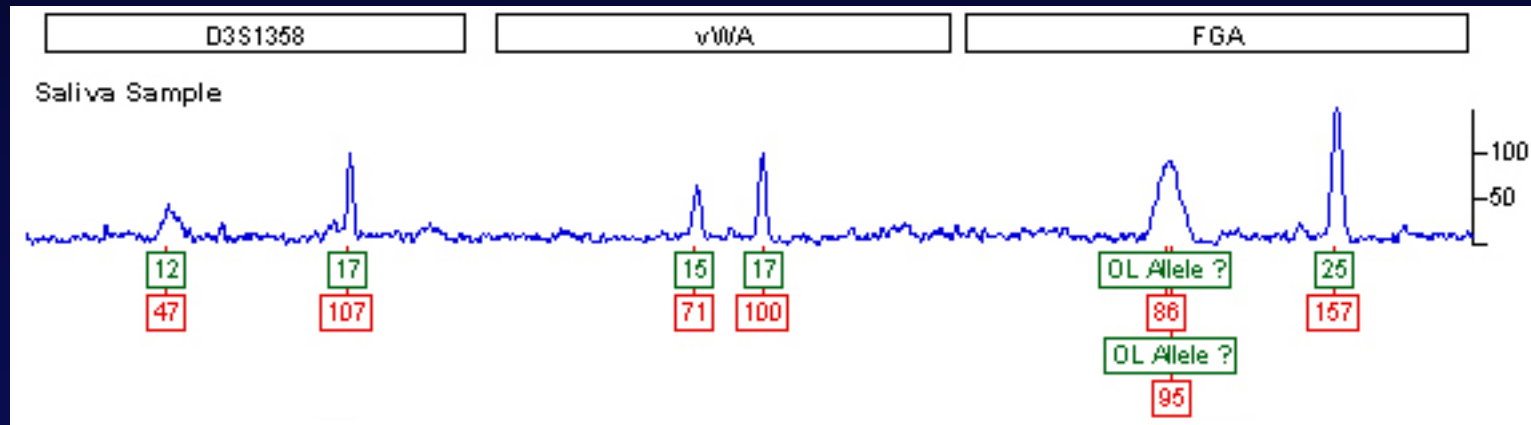


Can "Harry" be excluded?

| <u>Suspect</u> | <u>D3</u> | <u>vWA</u> | <u>FGA</u> |
|----------------|-----------|------------|------------|
| Tom            | 17, 17    | 15, 17     | 25, 25     |
| Dick           | 12, 17    | 15, 17     | 20, 25     |
| Harry          | 14, 17    | 15, 17     | 20, 25     |

No -- the 14 allele at D3 may be missing due to "allelic drop out"; FGA blob masks the 20 allele.

# Opportunities for subjective interpretation?



Can "Sally" be excluded?

| <u>Suspect</u> | <u>D3</u>     | <u>vWA</u>    | <u>FGA</u>    |
|----------------|---------------|---------------|---------------|
| Tom            | 17, 17        | 15, 17        | 25, 25        |
| Dick           | 12, 17        | 15, 17        | 20, 25        |
| Harry          | 14, 17        | 15, 17        | 20, 25        |
| <b>Sally</b>   | <b>12, 17</b> | <b>15, 15</b> | <b>20, 22</b> |

No -- there must be a second contributor;  
degradation explains the "missing" FGA allele.

# What can be done to make DNA testing more objective?

- Distinguish between signal and noise
- Deducing the number of contributors to mixtures
- Accounting for relatives
- Be mindful of the potential for human error

# Where do peak height thresholds come from (originally)?

- Applied Biosystems validation study of 1998
- Wallin et al., 1998, "TWGDAM validation of the AmpFISTR blue PCR Amplification kit for forensic casework analysis." *JFS* 43:854-870.

# Where do peak height thresholds come from (originally)?

PCR products were examined on both the 377 DNA Sequencer and the 310 Genetic Analyzer. The results of 0.25 to 1.0 ng were clearly typable with peak heights of approximately 150 RFU and greater (data not shown). At 0.125 ng and less, the peak heights in both samples were not significantly above the background ( $< 150$  RFU) or were undetectable. At 0.0313 ng specifically, peaks were extremely low or undetectable, and thus, DNA quantities as low as approximately 35 pg did not produce a typable result. Based on these results, we employed a peak height threshold of 150 RFU, below which peaks were interpreted with caution. Laboratories should determine a minimum peak height threshold for their instruments using high quality, single source genomic DNA samples which provides them with the desired sensitivity while not allowing for detection of low copy DNA. This is particularly important as the overall sensitivity of the assay may vary between laboratories.

# Where do peak height thresholds come from?

- “Conservative” thresholds established during validation studies
- Eliminate noise (even at the cost of eliminating signal)
- Can arbitrarily remove legitimate signal
- Contributions to noise vary over time (e.g. polymer and capillary age/condition)
- Analytical chemists use LOD and LOQ



# Signal Measure

Measured signal (In Volts/RFUS/etc)

# Saturation



## Quantification limit

$$\mu_b + 3\sigma_b$$

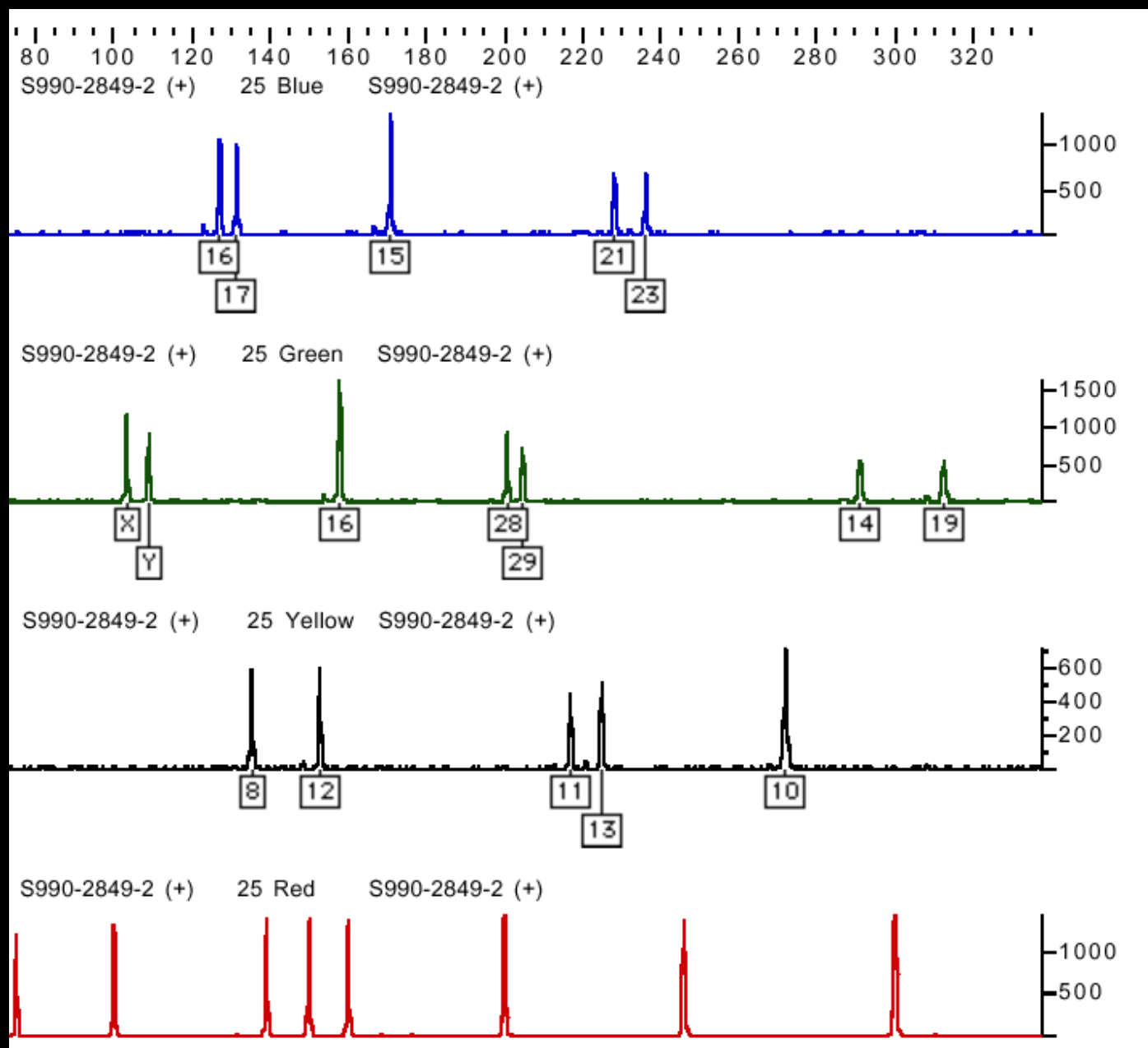
## Detection limit

 $\mu_b$ 

Mean background  
Signal

0

# Many opportunities to measure baseline



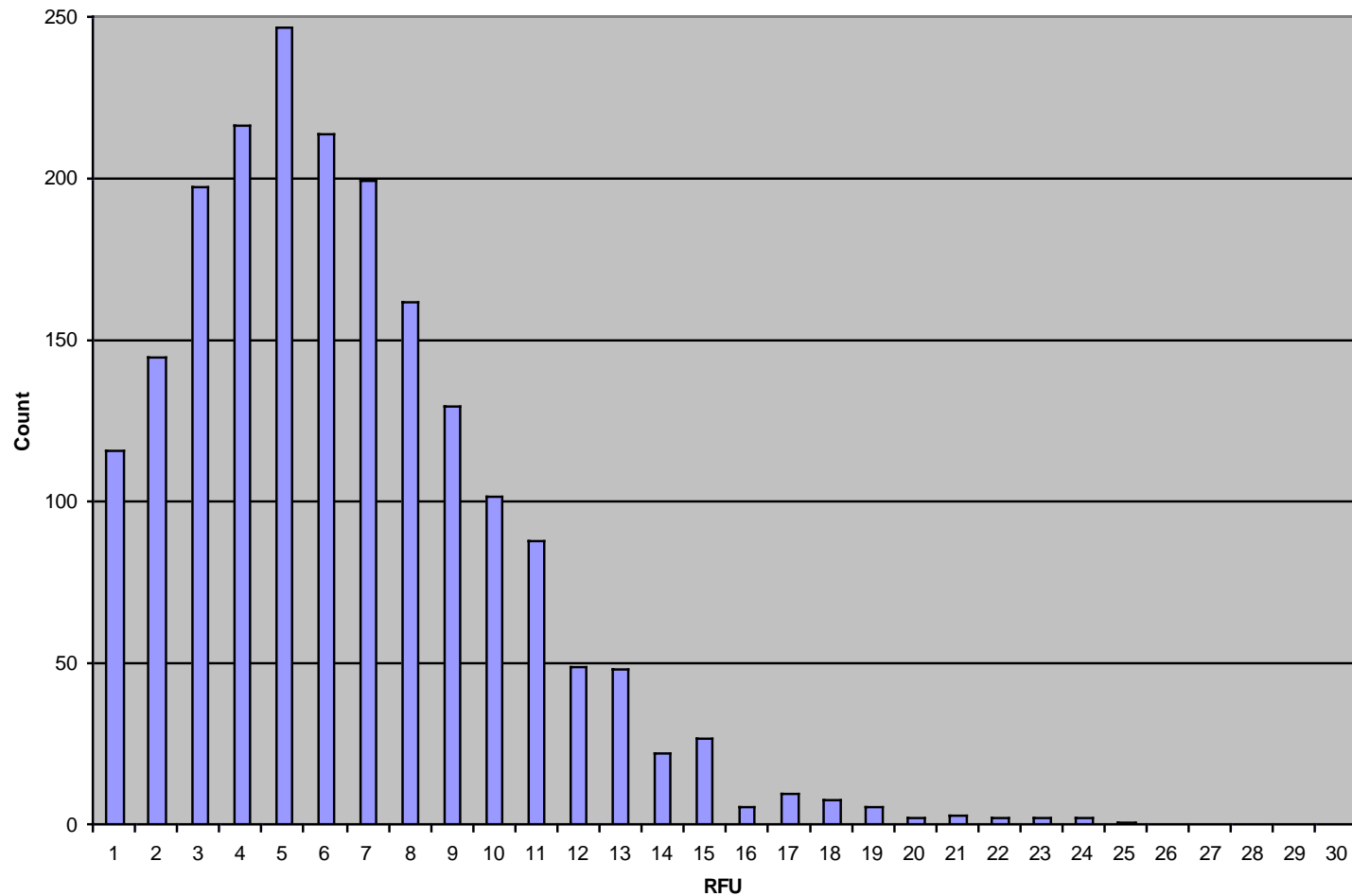
# Measurement of baseline in control samples:

- Negative controls: 5,932 data collection points (DCPs) per run ( $\sigma = 131$  DCPs)
- Reagent blanks: 5,946 DCPs per run ( $\sigma = 87$  DCPs)
- Positive controls: 2,415 DCP per run ( $\sigma = 198$  DCPs)

# Measurement of baseline in control samples:

- Negative controls: 5,932 data collection points (DCPs) per run ( $\sigma = 131$  DCPs)
- Reagent blanks: 5,946 DCPs per run ( $\sigma = 87$  DCPs)
- Positive controls: 2,415 DCP per run ( $\sigma = 198$  DCPs)
- DCP regions corresponding to size standards and 9947A peaks (plus and minus 55 DCPs to account for stutter in positive controls) were masked in all colors

# RFU levels at all non-masked data collection points

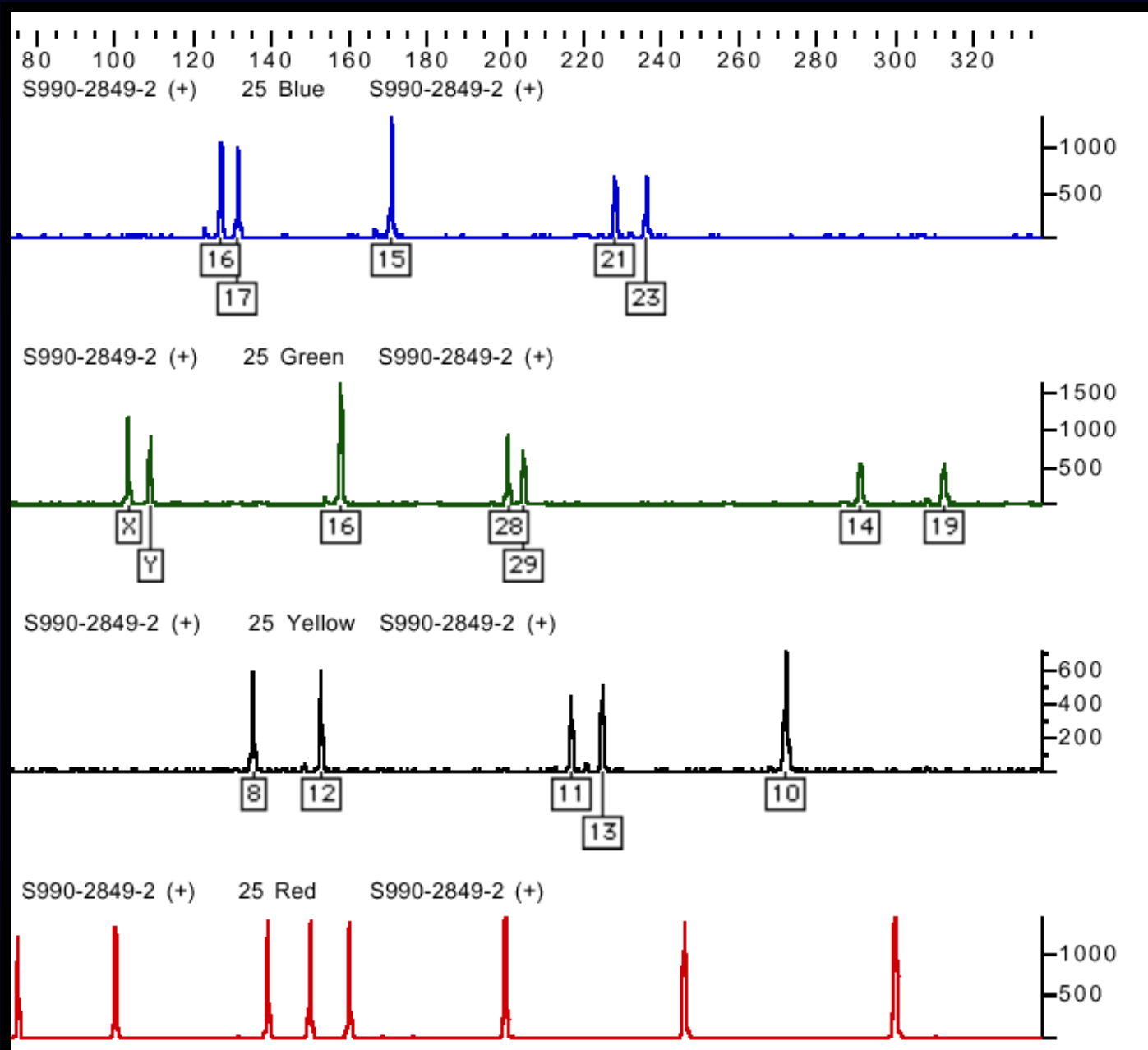


# Variation in baseline noise levels

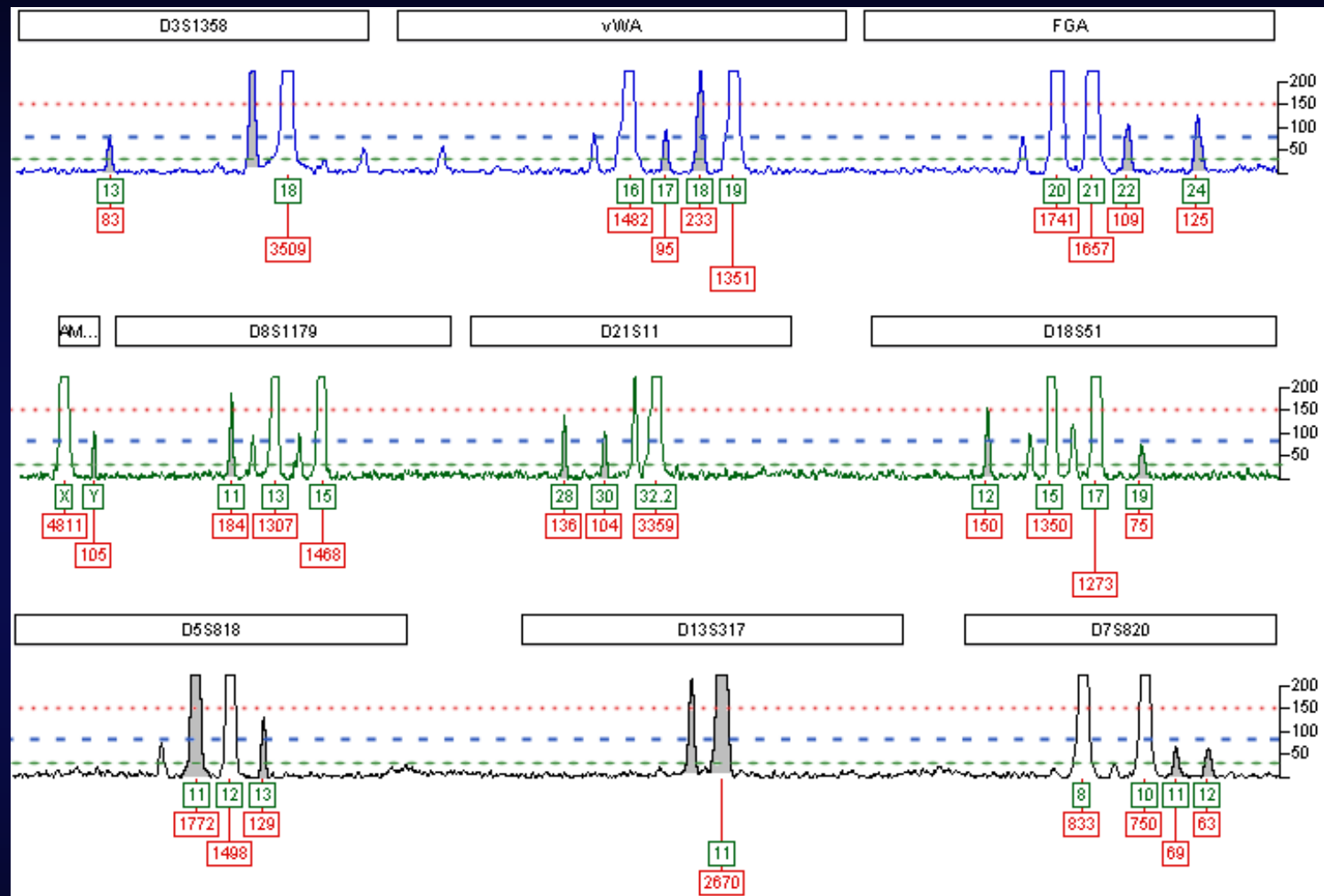
| Positive Control            |         | $\mu_b$ | $\sigma_b$ | $\mu_b + 3\sigma_b$ | $\mu_b + 10\sigma_b$ |
|-----------------------------|---------|---------|------------|---------------------|----------------------|
|                             | Maximum | 6.7     | 6.9        | 27.4                | 75.7                 |
|                             | Average | 5.0     | 3.7        | 16.1                | 42.0                 |
|                             | Minimum | 3.7     | 2.4        | 10.9                | 27.7                 |
| Negative Control            |         | $\mu_b$ | $\sigma_b$ | $\mu_b + 3\sigma_b$ | $\mu_b + 10\sigma_b$ |
|                             | Maximum | 13.4    | 13.2       | 53.0                | 145.4                |
|                             | Average | 5.4     | 3.9        | 17.1                | 44.4                 |
|                             | Minimum | 4.0     | 2.6        | 11.8                | 30.0                 |
| Reagent Blank               |         | $\mu_b$ | $\sigma_b$ | $\mu_b + 3\sigma_b$ | $\mu_b + 10\sigma_b$ |
|                             | Maximum | 6.5     | 11.0       | 39.5                | 116.5                |
|                             | Average | 5.3     | 4.0        | 17.3                | 45.3                 |
|                             | Minimum | 4.0     | 2.6        | 11.8                | 30.0                 |
| All three controls averaged |         | $\mu_b$ | $\sigma_b$ | $\mu_b + 3\sigma_b$ | $\mu_b + 10\sigma_b$ |
|                             | Maximum | 7.1     | 7.3        | 29.0                | 80.1                 |
|                             | Average | 5.2     | 3.9        | 16.9                | 44.2                 |
|                             | Minimum | 3.9     | 2.5        | 11.4                | 28.9                 |

Average ( $\bar{\mu}_b$ ) and standard deviation ( $\bar{\sigma}_b$ ) values with corresponding LODs and LOQs from positive, negative and reagent blank controls in 50 different runs. BatchExtract: <ftp://ftp.ncbi.nlm.nih.gov/pub/forensics/>

Doesn't someone either match or not?



# Lines in the sand: a two-person mix?



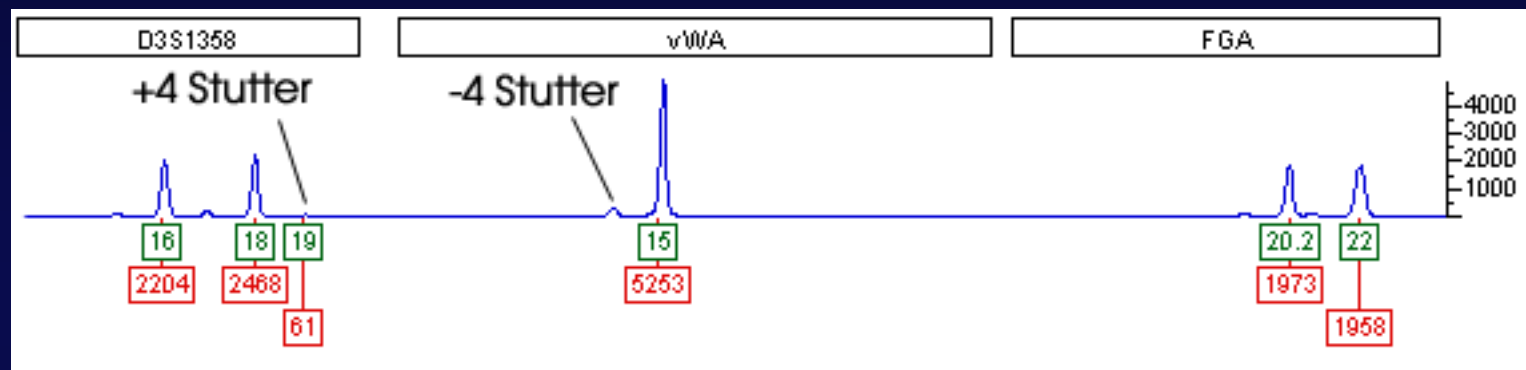
Two reference samples in a 1:10 ratio (male:female). Three different thresholds are shown: 150 RFU (red); LOQ at 77 RFU (blue); and LOD at 29 RFU (green).



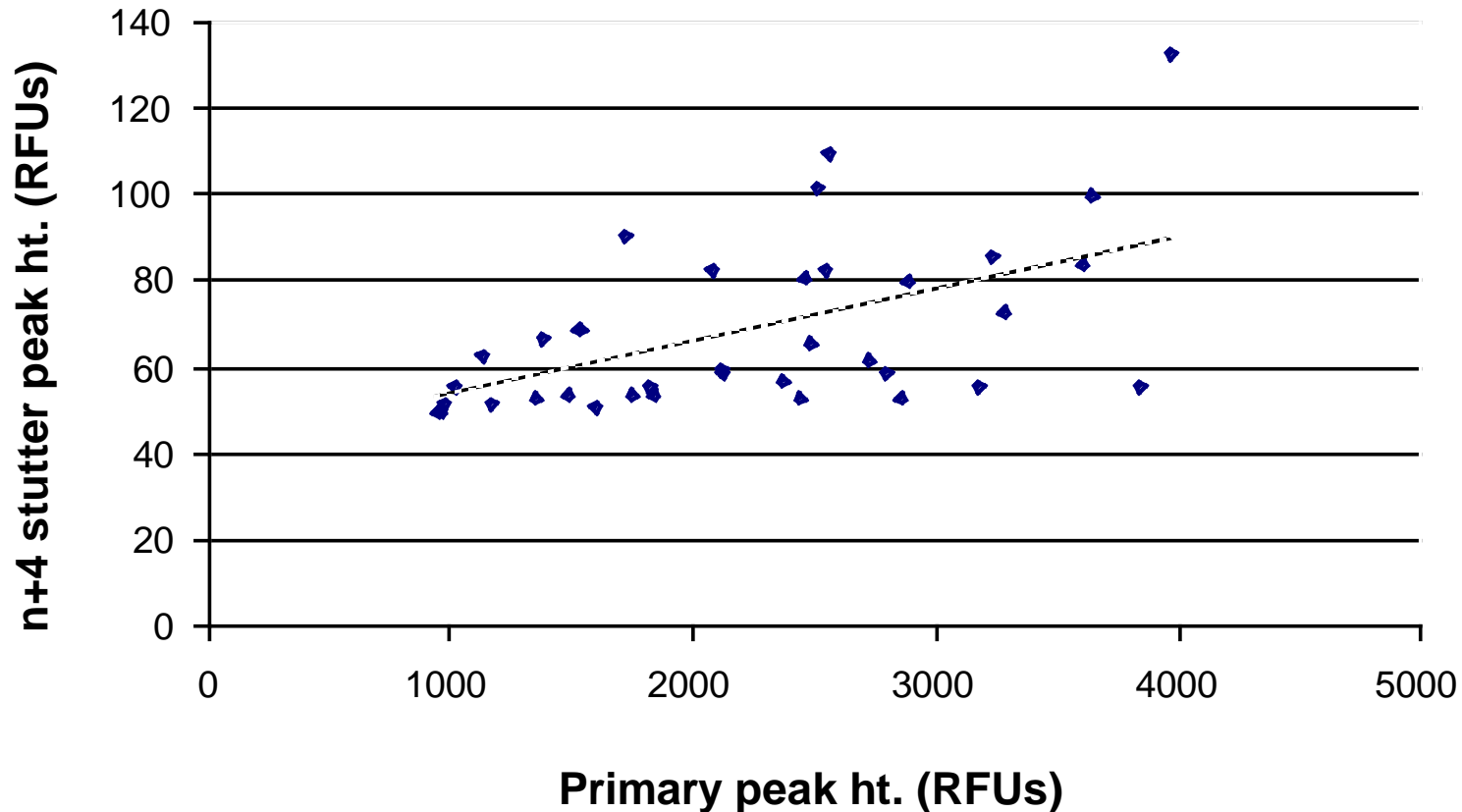
Not all signal comes from DNA  
associated with an evidence sample

- Stutter peaks
- Pull-up (bleed through)
- Spikes and blobs

# Stutter peaks

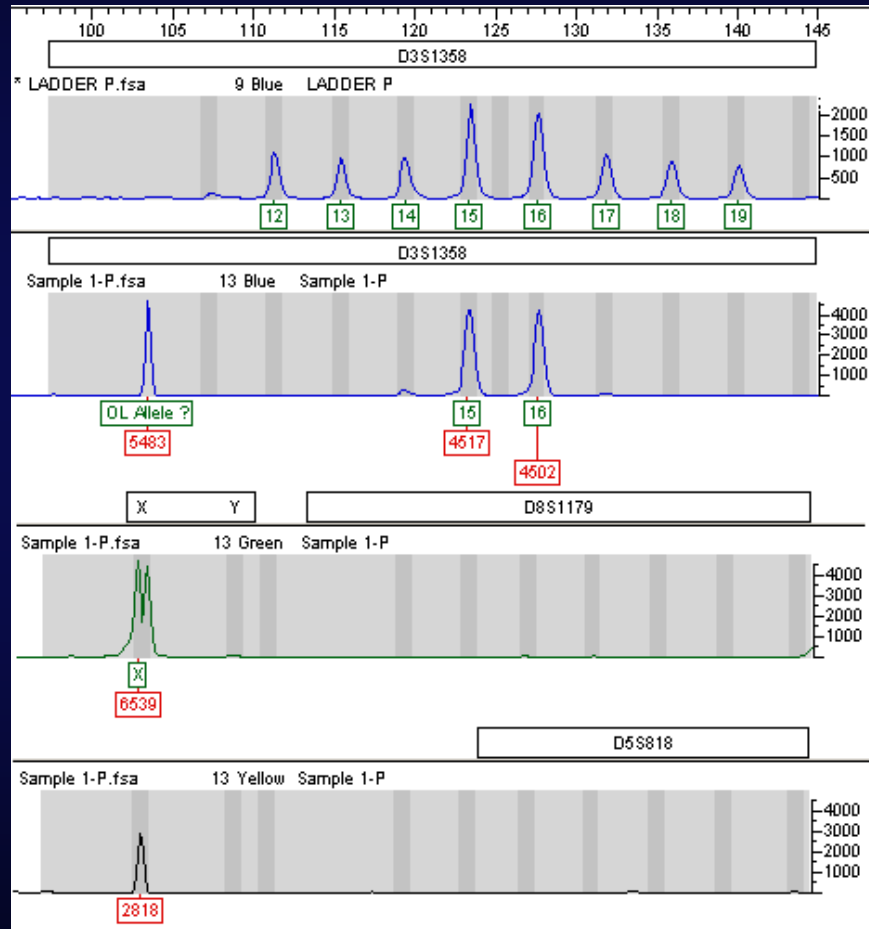


# The reality of n+4 stutter

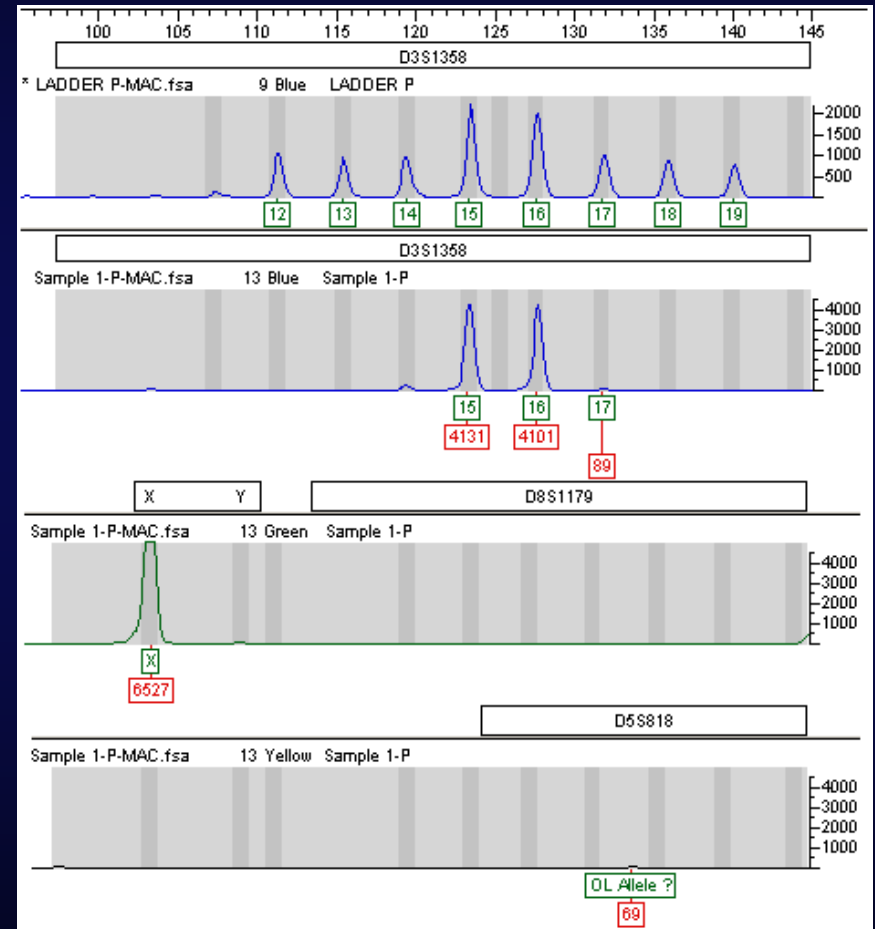


Primary peak height vs. n+4 stutter peak height. Evaluation of 37 data points,  $R^2=0.293$ ,  $p=0.0005$ . From 224 reference samples in 52 different cases. A filter of 5.9% would be conservative. Rowland and Krane, accepted with revision by *JFS*.

# Pull-up (and software differences)

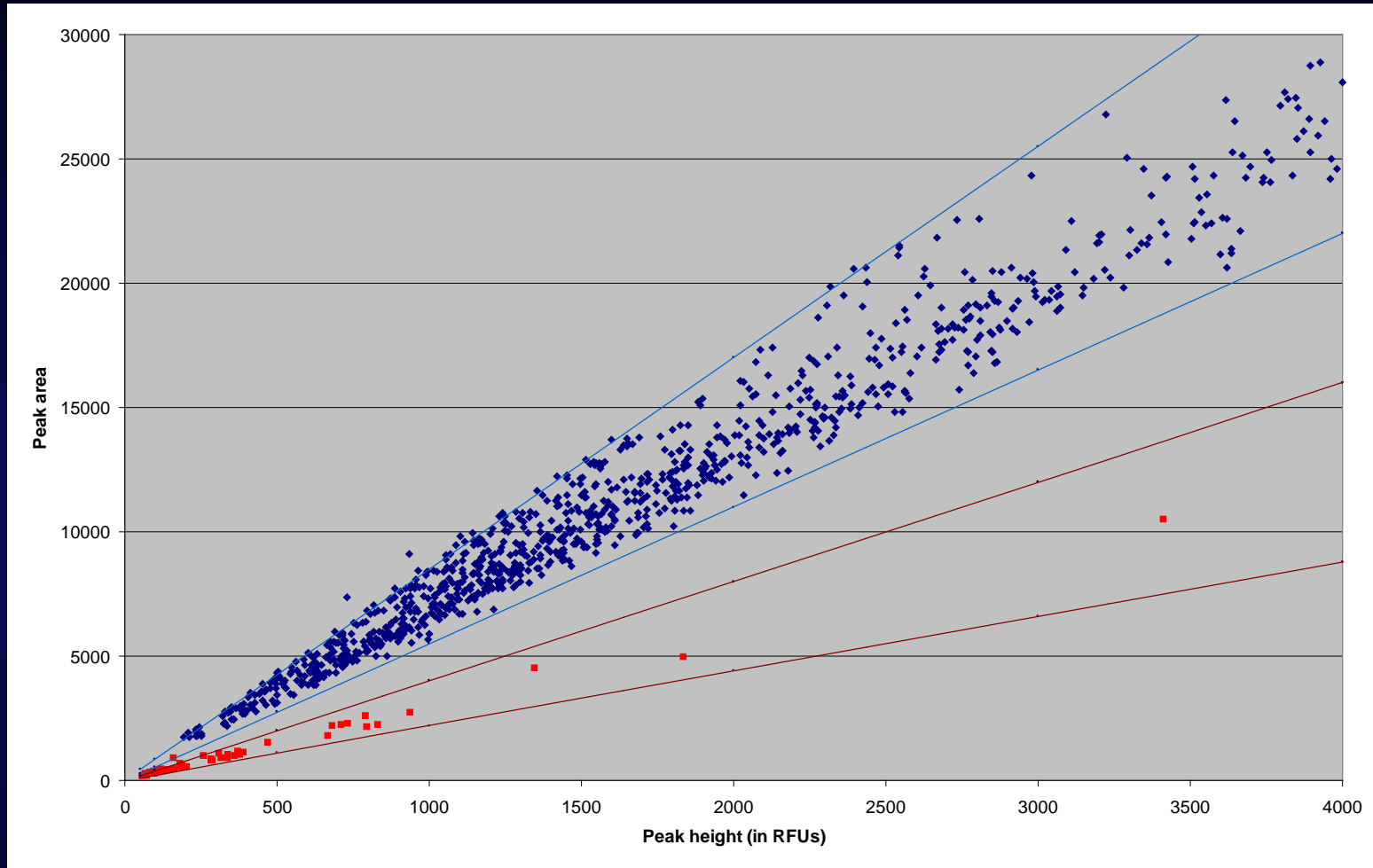


Advanced



Classic

# Spikes

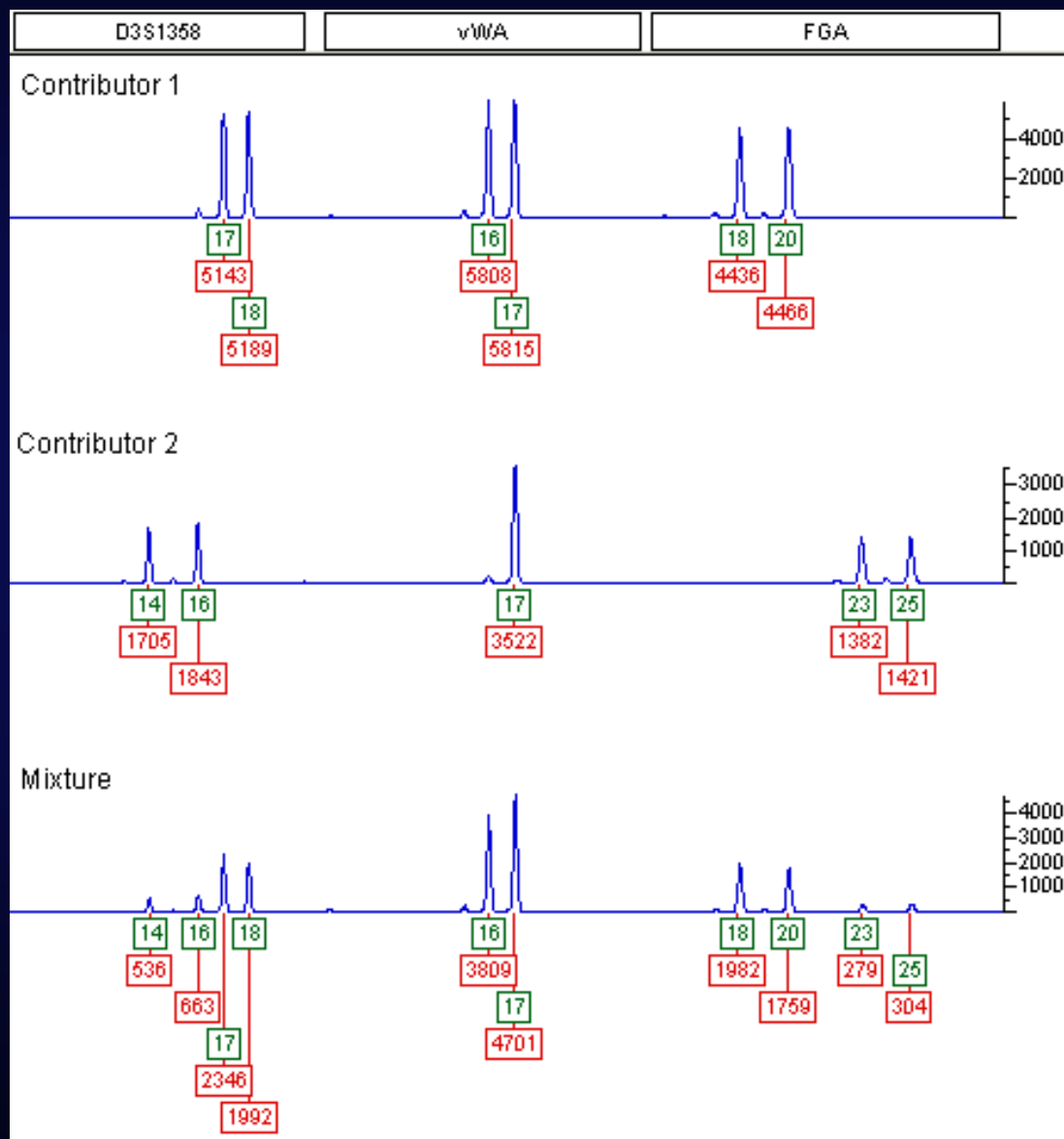


- 89 samples (references, pos controls, neg controls)
- 1010 "good" peaks
- 55 peaks associated with 24 spike events
- 95% boundaries shown

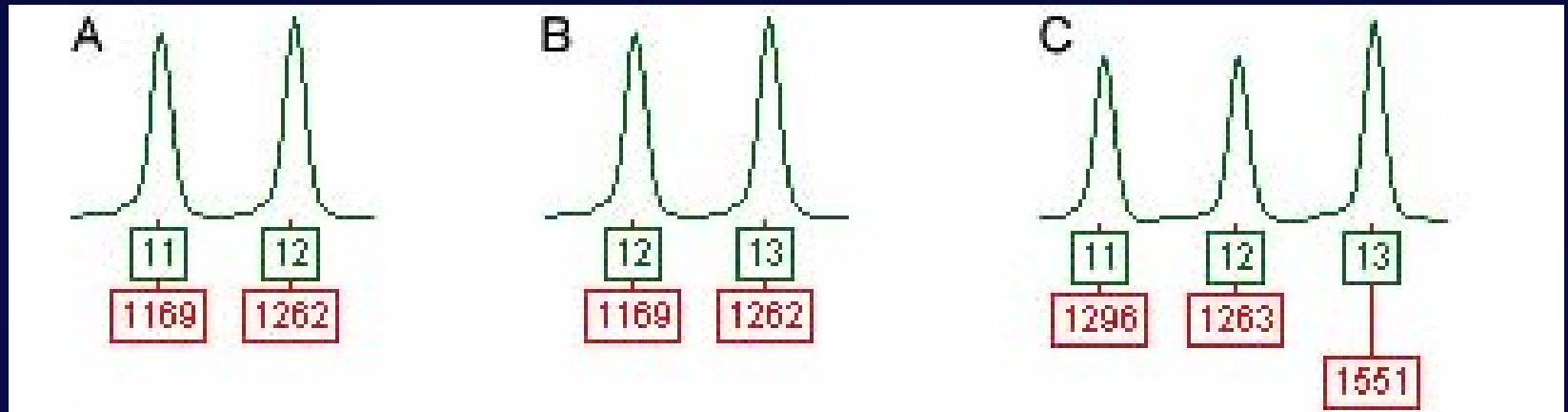
# What can be done to make DNA testing more objective?

- Distinguish between signal and noise
- Deducing the number of contributors to mixtures
- Accounting for relatives
- Be mindful of the potential for human error

# Mixed DNA samples



# Mixed DNA samples





# Suitable profiles for empirical mixing

- 959 complete 13-locus (CODIS-loci) STR genotypes used by the FBI for the purpose of allele frequency databases
- Includes: Bahamians (153); Trinidadians (76); US African Americans (177); Southwest Hispanics (202); Jamaicans (157); and US Caucasians (194)
- Available on-line at:  
<http://www.fbi.gov/hq/lab/fsc/backissu/july1999/dnaloci>
- Analyzed for Hardy-Weinberg equilibrium but no mention of possibility of relatives

# How many contributors to a mixture? If analysts can discard a locus?

| Maximum # of alleles observed in a 3-person mixture | # of occurrences          | Percent of cases |
|---|---------------------------|------------------|
| 2   | 0                         | 0.00             |
| 3   | 3,398<br>78               | 0.00<br>0.00     |
| 4   | 26,788,540<br>4,967,034   | 18.28<br>3.39    |
| 5   | 112,469,398<br>93,037,010 | 76.75<br>63.49   |
| 6   | 7,274,823<br>48,532,037   | 4.96<br>33.12    |

There are 146,536,159 possible different 3-person mixtures of the 959 individuals in the FBI database (Paoletti et al., November 2005 *JFS*).

# How many contributors to a mixture? If analysts can discard a locus?

| Maximum # of alleles observed in a 3-person mixture | # of occurrences         | Percent of cases |
|---|--------------------------|------------------|
| 2   | 0                        | 0.00             |
| 3   | 8,151<br>310             | 0.02<br>0.00     |
| 4   | 11,526,219<br>2,498,139  | 25.53<br>5.53    |
| 5   | 32,078,976<br>29,938,777 | 71.07<br>66.32   |
| 6   | 1,526,550<br>12,702,670  | 3.38<br>28.14    |

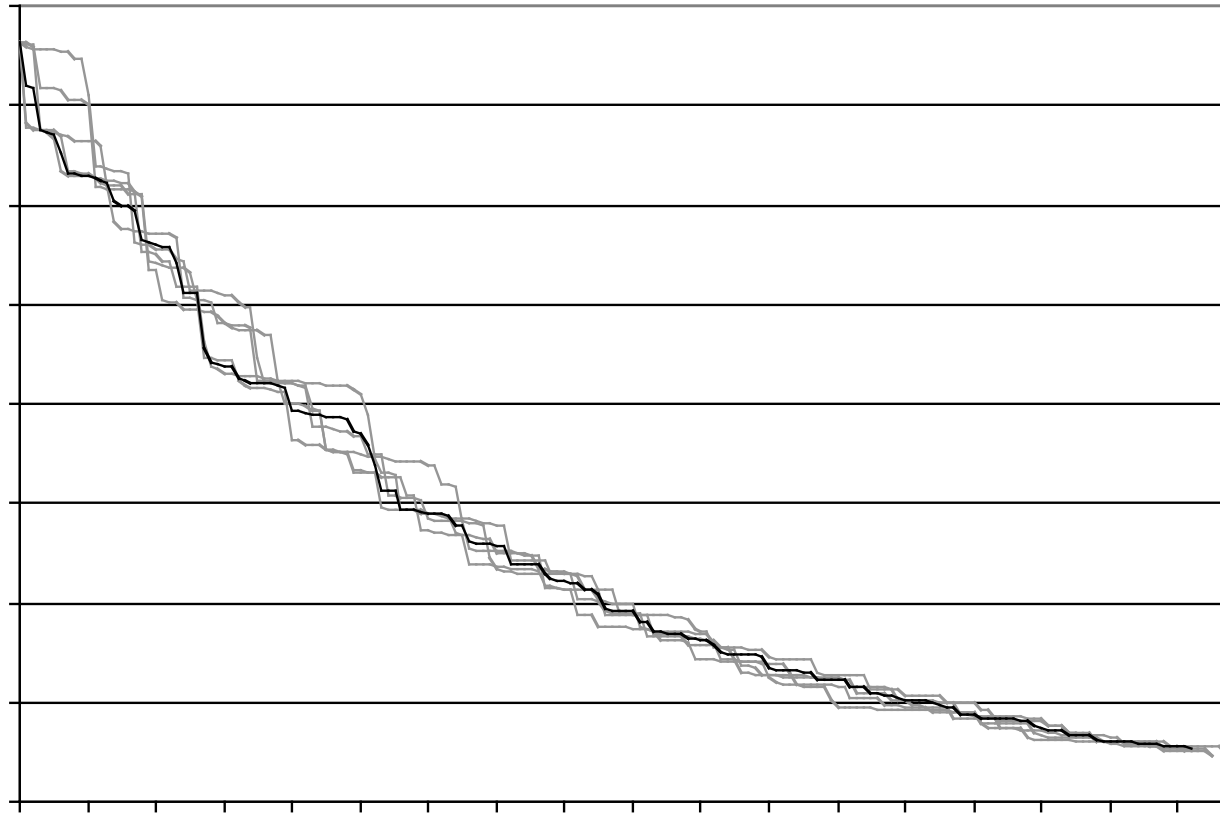
There are 45,139,896 possible different 3-person mixtures of the 648 individuals in the MN BCI database (genotyped at only 12 loci).

# How many contributors to a mixture?

| Maximum # of<br>alleles observed in<br>a 4-person mixture | # of occurrences | Percent of cases |
|---|------------------|------------------|
| 4   | 13,480           | 0.02             |
| 5   | 8,596,320        | 15.03            |
| 6   | 35,068,040       | 61.30            |
| 7   | 12,637,101       | 22.09            |
| 8   | 896,435          | 1.57             |

There are 57,211,376 possible different 4-way mixtures of the 194 individuals in the FBI Caucasian database (Paoletti et al., November 2005 *JFS*). (35,022,142,001 4-person mixtures with 959 individuals.)

# Does testing more loci help?

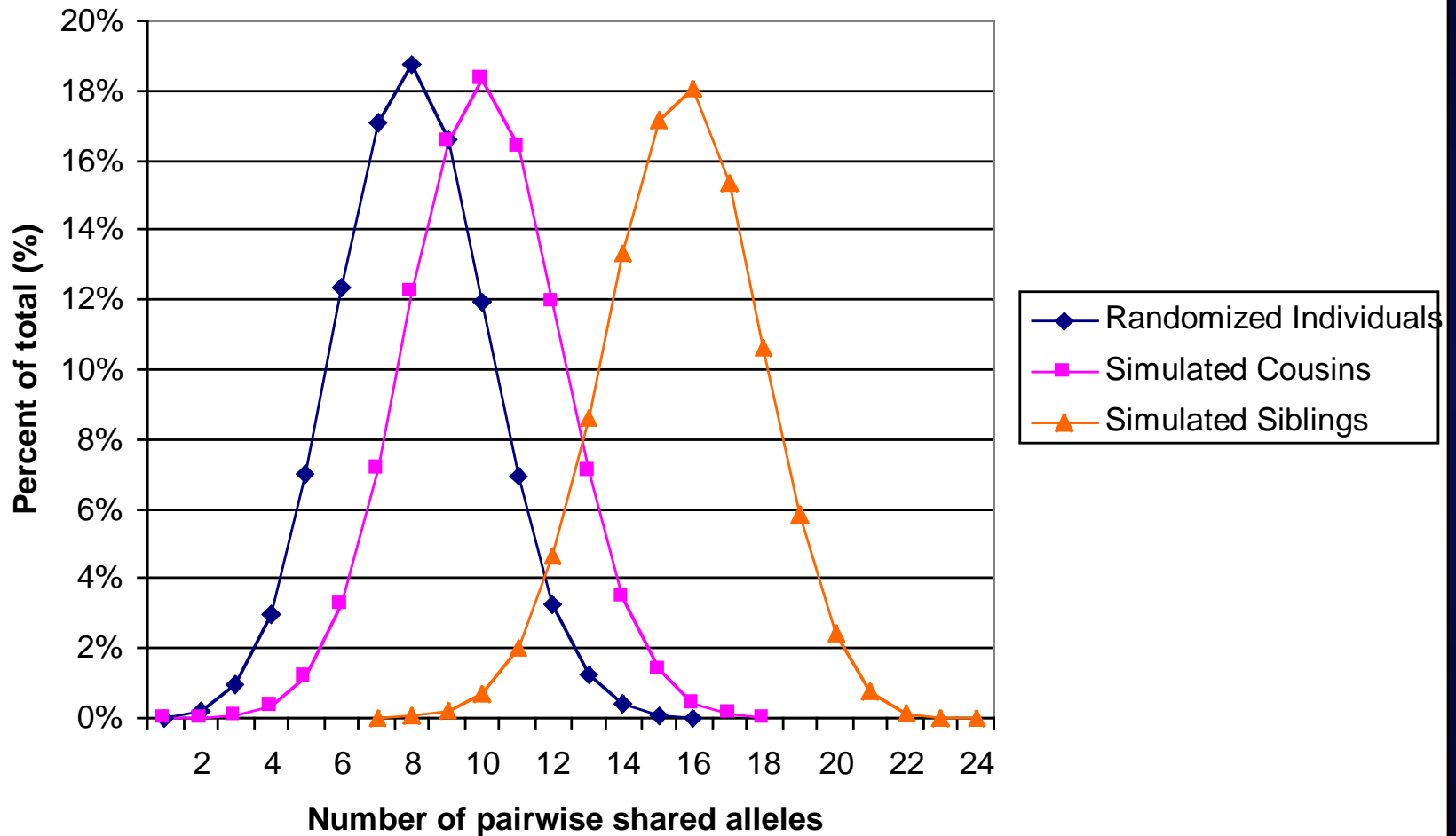


Five simulations are shown with each data point representing 57,211,376 4-person mixtures (average shown in black). (Paoletti et al., November 2005 *JFS*). Mischaracterization rate of 76.34% for original 13 loci.

# What contributes to overlapping alleles between individuals?

- Identity by state
  - many loci have a small number of detectable alleles (only 6 for TPOX and 7 for D13, D5, D3 and TH01)
  - some alleles at some loci are relatively common
- Identity by descent
  - relatives are more likely to share alleles than unrelated individuals
  - perfect 13 locus matches between siblings occur at an average rate of 3.0 per 459,361 sibling pairs

# Allele sharing between individuals



## Allele sharing in databases

- Original FBI dataset's mischaracterization rate for 3-person mixtures (3.39%) is more than two  $\sigma$  above the average observed in five sets of randomized individuals
- Original FBI dataset has more shared allele counts above 19 than five sets of randomized individuals (3 vs. an average of 1.4)



# Final thoughts on mixed samples

- Maximum allele count by itself is not a reliable predictor of the number of contributors to mixed forensic DNA samples.
- Simply reporting that a sample “arises from two or more individuals” is reasonable and appropriate.
- Analysts should exercise great caution when invoking discretion.
- Excess allele sharing observed in the FBI allele frequency database is most easily explained by the presence of relatives in that database.

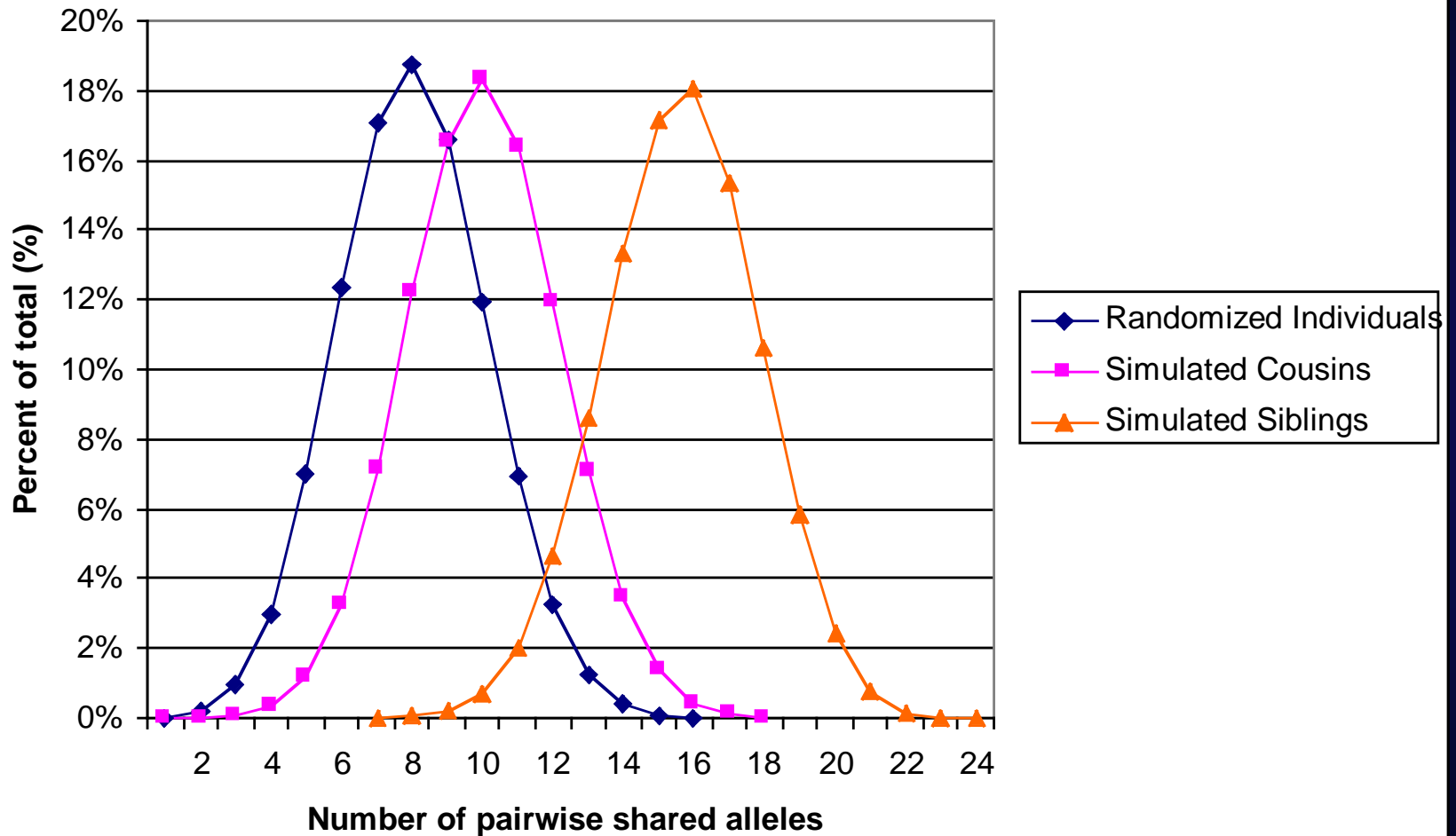
# Familial searching

- Database search yields a close but imperfect DNA match
- Can suggest a relative is the true perpetrator
- Great Britain performs them routinely
- Reluctance to perform them in US since 1992 NRC report
- Current CODIS software cannot perform effective searches

# Three approaches to familial searches

- Search for rare alleles (inefficient)
- Count matching alleles (arbitrary)
- Likelihood ratios with kinship analyses

# Pair-wise similarity distributions



# Is the true DNA match a relative or a random individual?

- Given a closely matching profile, who is more likely to match, a relative or a randomly chosen, unrelated individual?
- Use a likelihood ratio

$$LR = \frac{P(E \mid \textit{relative})}{P(E \mid \textit{random})}$$

# Is the true DNA match a relative or a random individual?

- What is the likelihood that a relative of a single initial suspect would match the evidence sample perfectly?
- What is the likelihood that a single randomly chosen, unrelated individual would match the evidence sample perfectly?

$$LR = \frac{P(E \mid \textit{relative})}{P(E \mid \textit{random})}$$

# Probabilities of siblings matching at 0, 1 or 2 alleles

$$P(E | sib) = \begin{cases} \frac{P_a \cdot P_b \cdot HF}{4}, & \text{if } shared = 0 \\ \frac{P_b + P_a \cdot P_b \cdot HF}{4}, & \text{if } shared = 1 \\ \frac{1 + P_a + P_b + P_a \cdot P_b \cdot HF}{4}, & \text{if } shared = 2 \end{cases}$$

HF = 1 for homozygous loci and 2 for heterozygous loci;  
P<sub>a</sub> is the frequency of the allele shared by the evidence sample and the individual in a database.

# Probabilities of parent/child matching at 0, 1 or 2 alleles

$$P(E \mid \text{parent / child}) = \begin{cases} 0, & \text{if } \text{shared} = 0 \\ \frac{P_b}{2}, & \text{if } \text{shared} = 1 \\ \frac{P_a + P_b}{2}, & \text{if } \text{shared} = 2 \end{cases}$$

HF = 1 for homozygous loci and 2 for heterozygous loci;  
 $P_a$  is the frequency of the allele shared by the evidence sample and the individual in a database.



# Other familial relationships

Cousins:

$$P(E | \text{cousins}) = \begin{cases} \frac{6 \cdot P_a \cdot P_b \cdot HF}{8}, & \text{if } \text{shared} = 0 \\ \frac{P_b + 6 \cdot P_a \cdot P_b \cdot HF}{8}, & \text{if } \text{shared} = 1 \\ \frac{P_a + P_b + 6 \cdot P_a \cdot P_b \cdot HF}{8}, & \text{if } \text{shared} = 2 \end{cases}$$

Grandparent-grandchild;  
aunt/uncle-nephew-  
niece; half-siblings:

$$P(E | GG / AUNN / HS) = \begin{cases} \frac{2 \cdot P_a \cdot P_b \cdot HF}{4}, & \text{if } \text{shared} = 0 \\ \frac{P_b + 2 \cdot P_a \cdot P_b \cdot HF}{4}, & \text{if } \text{shared} = 1 \\ \frac{P_a + P_b + 2 \cdot P_a \cdot P_b \cdot HF}{4}, & \text{if } \text{shared} = 2 \end{cases}$$

HF = 1 for homozygous loci and 2 for heterozygous loci;  
 $P_a$  is the frequency of the allele shared by the evidence  
sample and the individual in a database.

# Familial search experiment

- Randomly pick related pair or unrelated pair from a synthetic database
- Choose one profile to be evidence and one profile to be initial suspect
- Test hypothesis:
  - $H_0$ : A relative is the source of the evidence
  - $H_A$ : An unrelated person is the source of the evidence

Paoletti, D., Doom, T., Raymer, M. and Krane, D. 2006. Assessing the implications for close relatives in the event of similar but non-matching DNA profiles. *Jurimetrics*, 46:161-175.

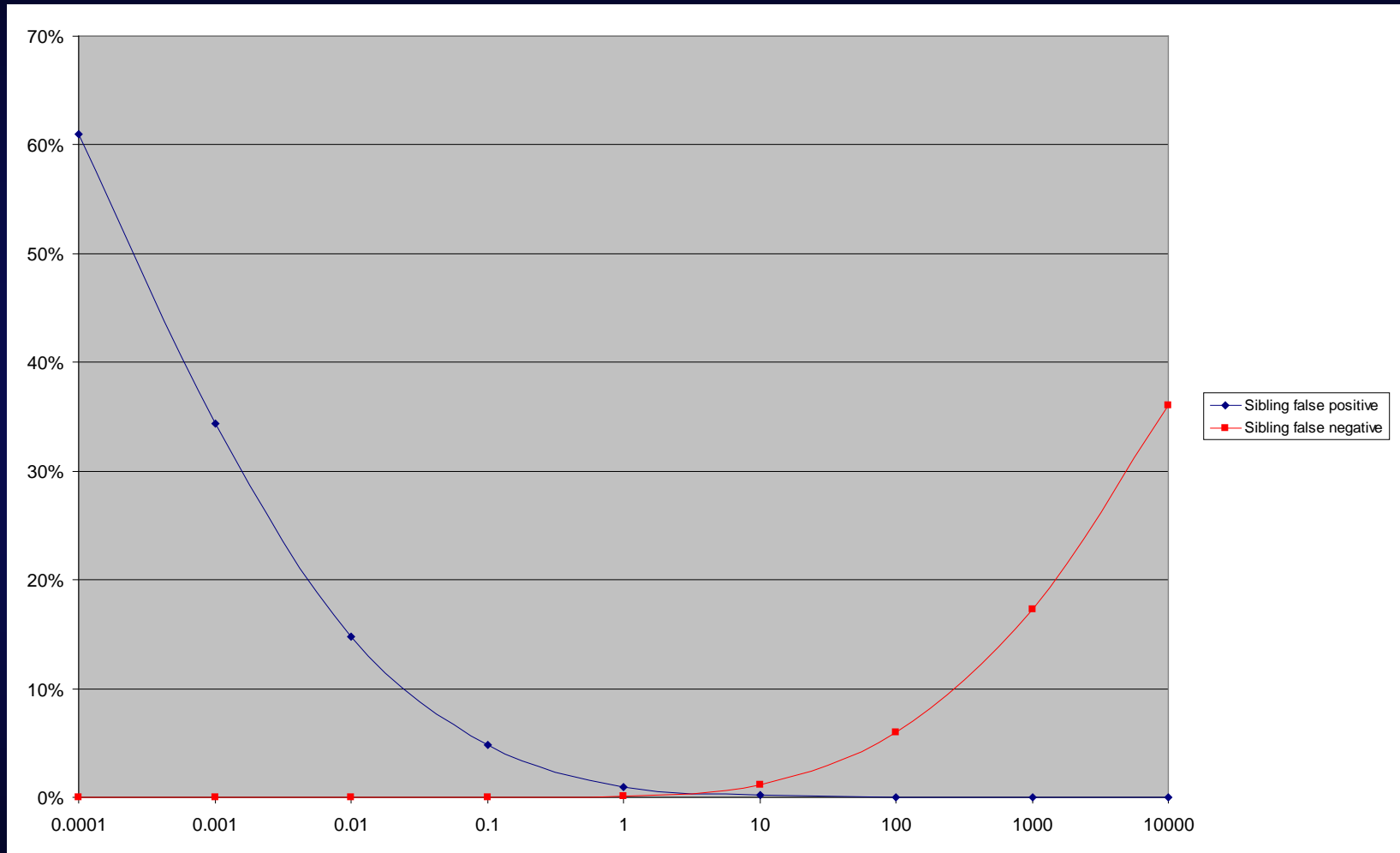
# Hypothesis testing using an LR threshold of 1 with prior odds of 1

|                 |                                    | True state                             |   |
|-----------------|------------------------------------|--|---|
|                 |                                    | Evidence from Unrelated individual     | Evidence from sibling                   |
| <b>Decision</b> | Evidence from unrelated individual | ~ 98%<br>[Correct decision]            | ~ 4%<br>[Type II error; false negative] |
|                 | Evidence from sibling              | ~ 2%<br>[Type I error; false positive] | ~ 96%<br>[Correct decision]             |

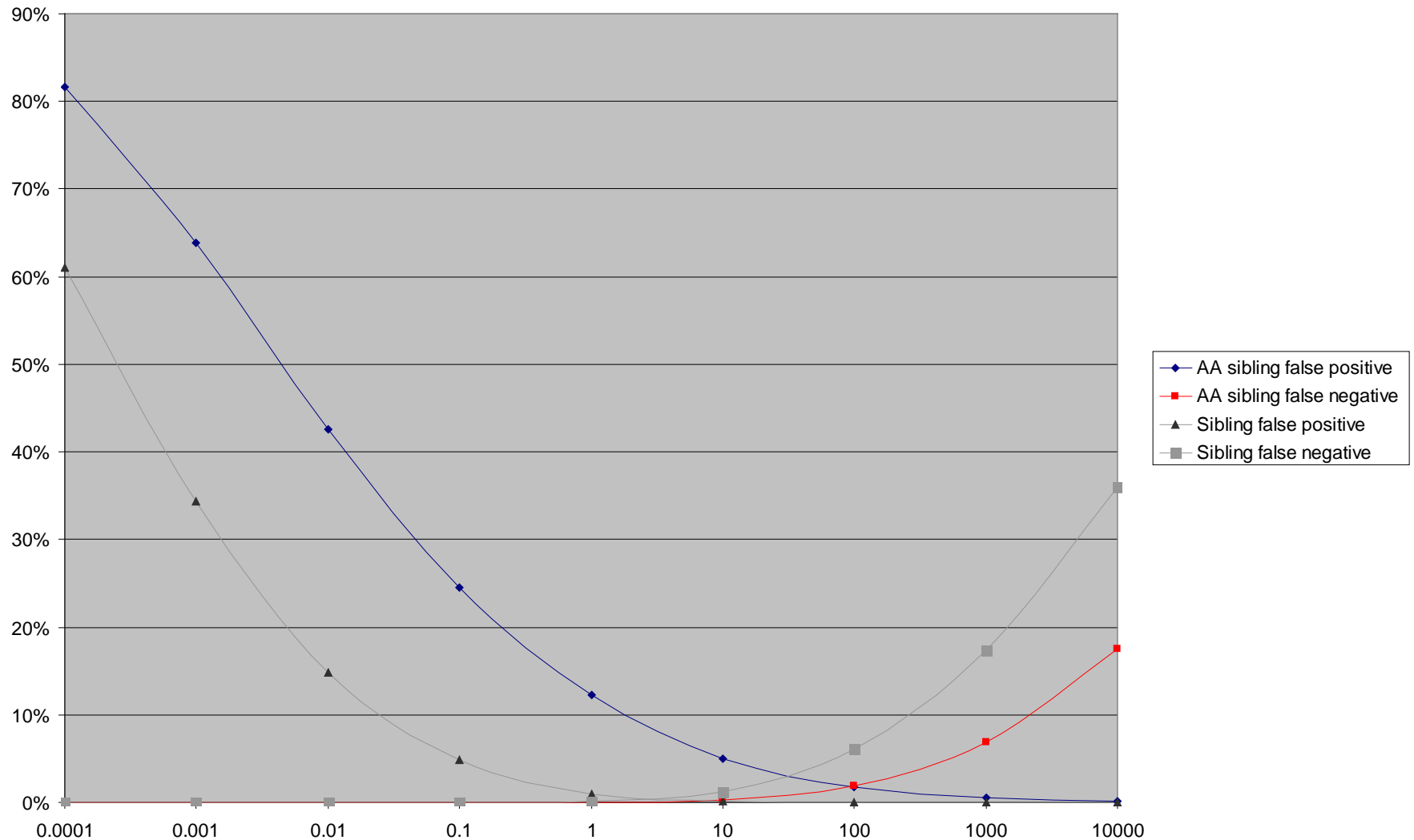
# Two types of errors

- False positives (Type I): an initial suspect's family is investigated even though an unrelated individual *is* the actual source of the evidence sample.
- False negatives (Type II): an initial suspect's family is *not* be investigated even though a relative really is the source of the evidence sample.
- A wide net (low LR threshold) catches more criminals but comes at the cost of more fruitless investigations.

# Type I and II errors with prior odds of 1



# Type I and II errors with prior odds of 1 and non-cognate allele frequencies



# Is the true DNA match a relative or a random individual?

- What is the likelihood that a close relative of a single initial suspect would match the evidence sample perfectly?
- What is the likelihood that a single randomly chosen, unrelated individual would match the evidence sample perfectly?

$$LR = \frac{P(E \mid \textit{relative})}{P(E \mid \textit{random})}$$

# Is the true DNA match a relative or a random individual?

- What is the likelihood that the source of the evidence sample was a relative of an initial suspect?

$$P(sib | E) = \frac{P(E | sib) \cdot P(sib)}{P(E | sib) \cdot P(sib) + P(E | random) \cdot P(random)}$$

Prior odds:

$$P(sib) = \frac{s}{popsize}$$

$$P(random) = \frac{popsize - s}{popsize}$$

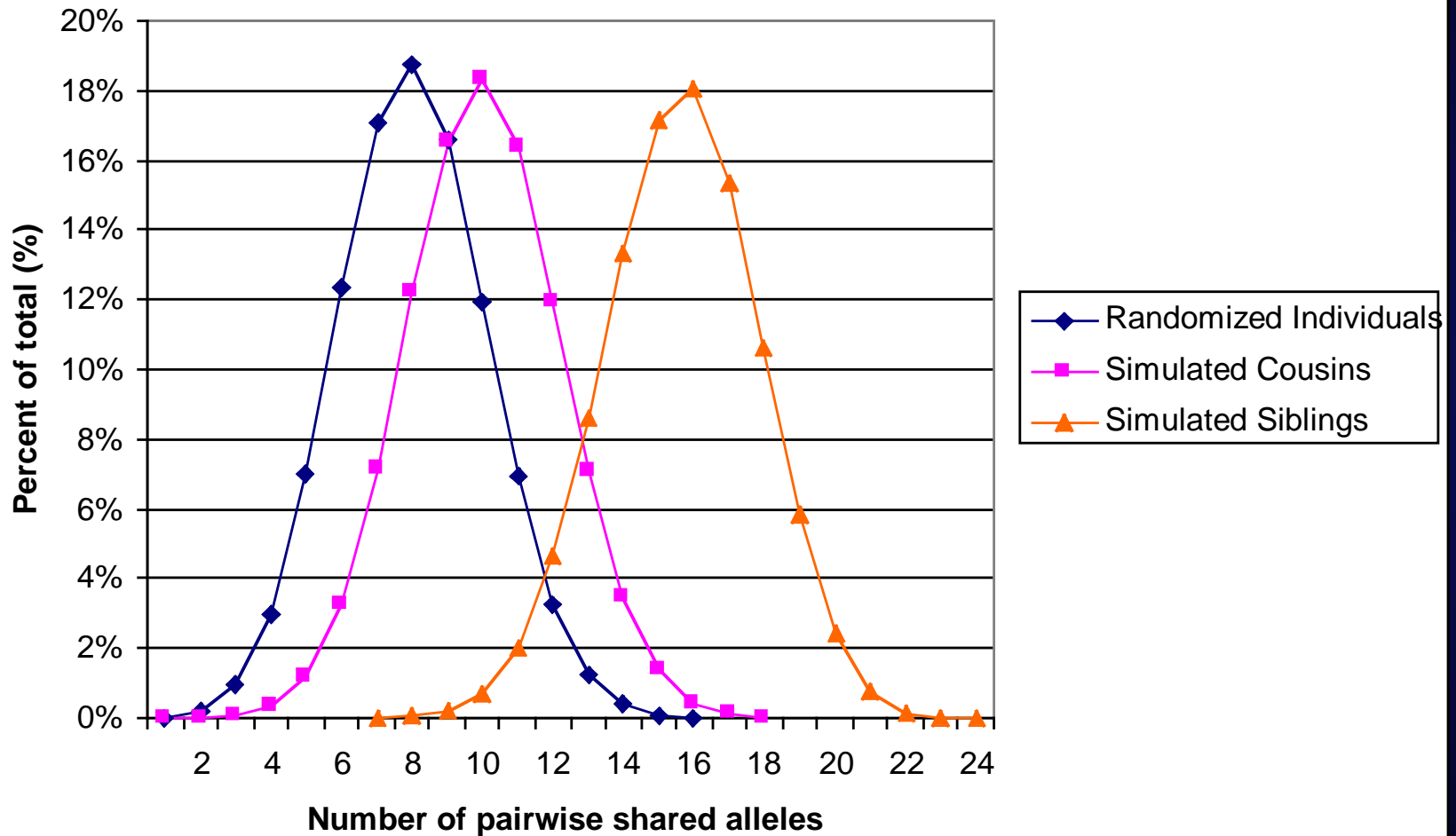


# Is the true DNA match a relative or a random individual?

- This more difficult question is ultimately governed by two considerations:
  - What is the size of the alternative suspect pool?
  - What is an acceptable rate of false positives?

$$LR = \frac{P(E \mid sib)}{P(E \mid random)}$$

# Pair-wise similarity distributions



# How well does an LR approach perform relative to alternatives?

- Low-stringency CODIS search identifies all 10,000 parent-child pairs (but only 1,183 sibling pairs and less than 3% of all other relationships and a high false positive rate)
- Moderate and high-stringency CODIS searches failed to identify any pairs for any relationship
- An allele count-threshold (set at 20 out of 30 alleles) identifies 4,233 siblings and 1,882 parent-child pairs (but fewer than 70 of any other relationship and with no false positives)

# How well does an LR approach perform relative to alternatives?

- LR set at 1 identifies > 99% of both sibling and parent-child pairs (with false positive rates of 0.01% and 0.1%, respectively)
- LR set at 10,000 identifies 64% of siblings and 56% of parent-child pairs (with no false positives)
- Use of non-cognate allele frequencies results in an increase in false positives and a decrease in true positives (that are largely offset by either a ceiling or consensus approach)

# How well does an LR approach perform relative to alternatives?

- LR set at 1 identifies > 78% of half-sibling, aunt-niece, and grandparent-grandchild pairs (with false positive rates at or below 9%)
- LR set at 1 identifies 58% of cousin pairs (with a 19% false positive rate)
- LR set at 10,000 identifies virtually no half-sibling, aunt-niece, grandparent-grandchild or cousin pairs (with no false positives)

# How well does an LR approach perform with mixed samples?

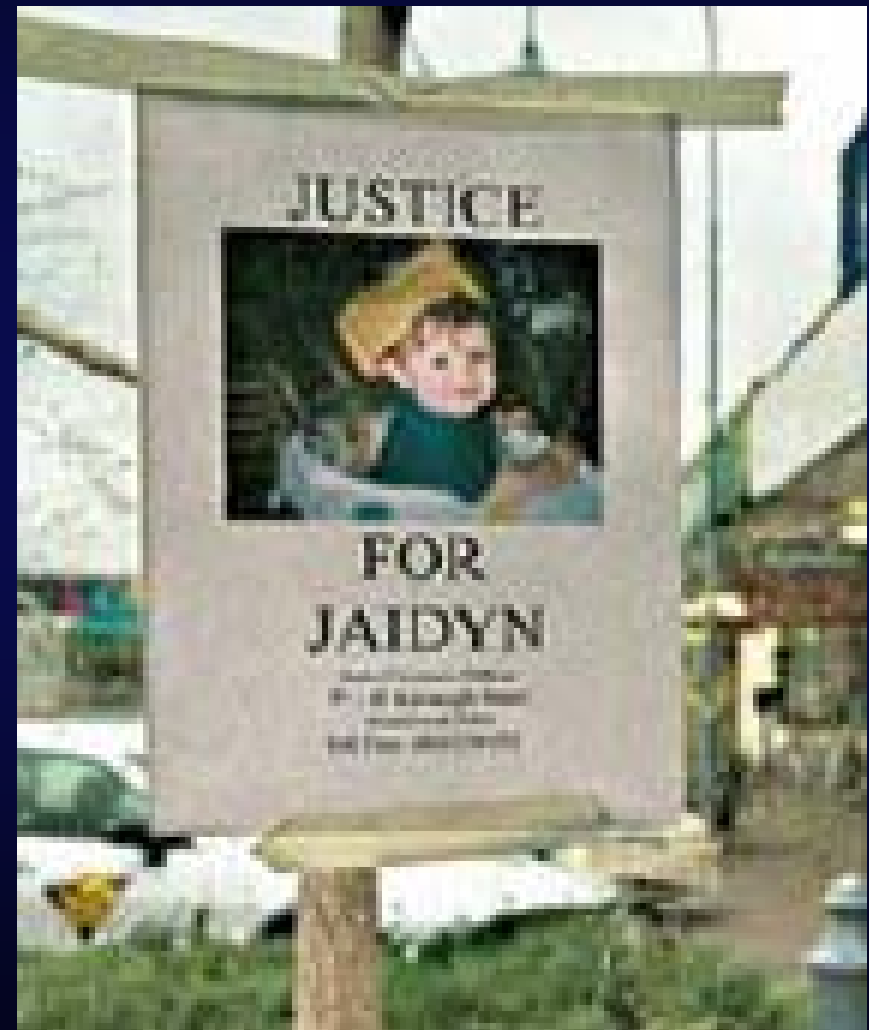
- LR set at 1 identifies >99% of both sibling and parent-child pairs even in 2- and 3-person mixtures (with false positive rates of 10% and 15%, and of 0.01% and 0.07%, respectively)
- LR set at 1 identifies >86% of half-sibling, aunt-niece, and grandparent-grandchild pairs in 2- and 3-person mixtures (with false positive rates lower than 22% and 30%, respectively)
- LR set at 1 identifies >74% of cousin pairs in 2- and 3-person mixtures (with false positive rates of 41% and 49%, respectively)

# What can be done to make DNA testing more objective?

- Distinguish between signal and noise
- Deducing the number of contributors to mixtures
- Accounting for relatives
- Be mindful of the potential for human error

# Victorian Coroner's inquest into the death of Jaidyn Leskie

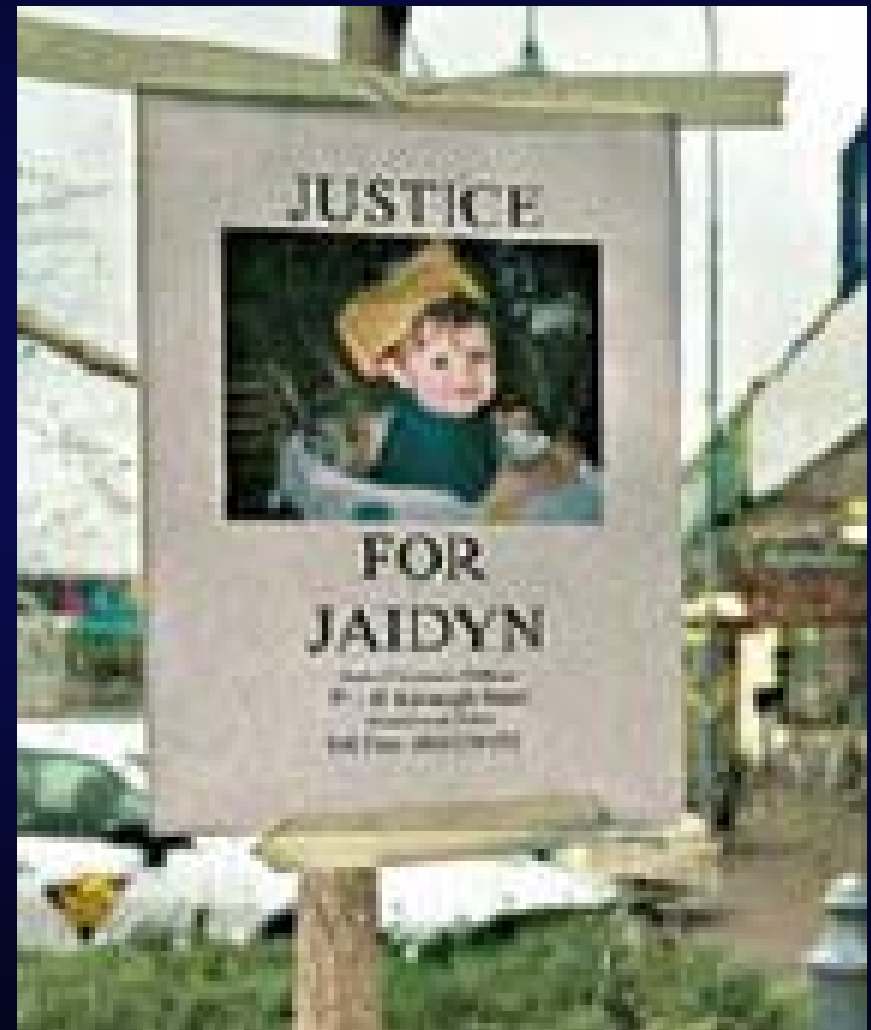
- Toddler disappears in bizarre circumstances: found dead six months later
- Mother's boy friend is tried and acquitted.
- Unknown female profile on clothing.
- Cold hit to a rape victim.
- RMP: 1 in 227 million.
- Lab claims "adventitious match."



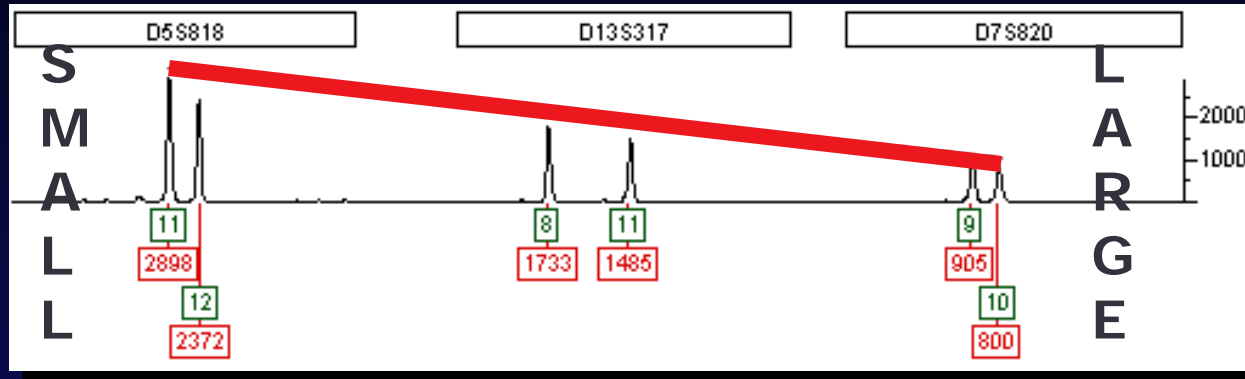


# Victorian Coroner's inquest into the death of Jaidyn Leskie

- Condom with rape victim's DNA was processed in the same lab 1 or 2 days prior to Leskie samples.
- Additional tests find matches at 5 to 7 more loci.
- Review of electronic data reveals low level contributions at even more loci.
- Degradation study further suggests contamination.



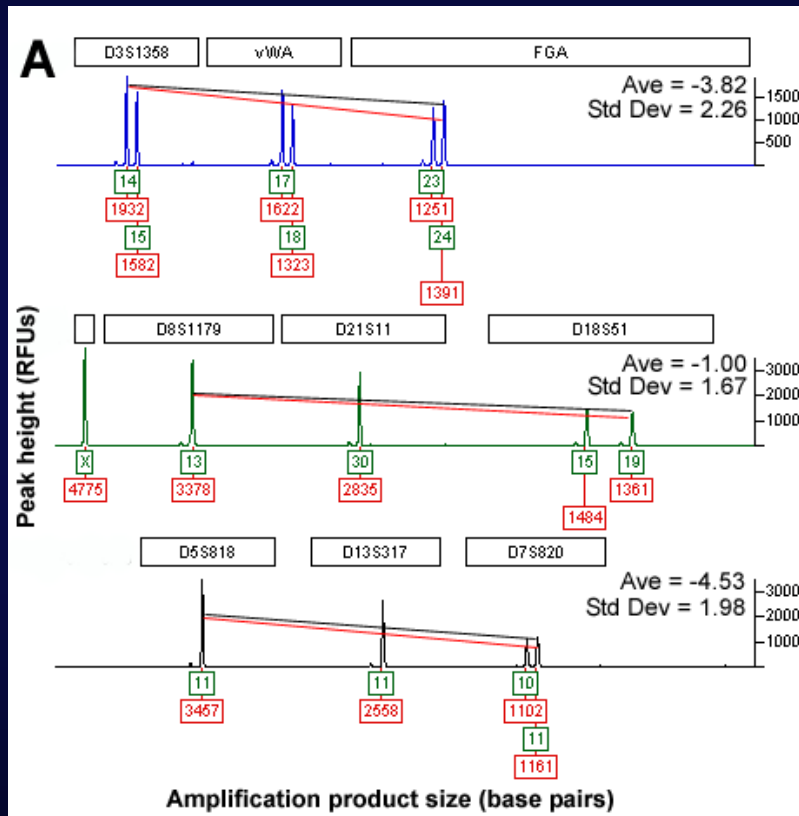
# Degradation, inhibition



- When biological samples are exposed to adverse environmental conditions, they can become degraded
  - **Warm, moist, sunlight, time**
- Degradation breaks the DNA at random
- Larger amplified regions are affected first
- Classic 'ski-slope' electropherogram
- Degradation and inhibition are unusual and noteworthy.

# Degradation, inhibition

## The Leskie Inquest, a practical application

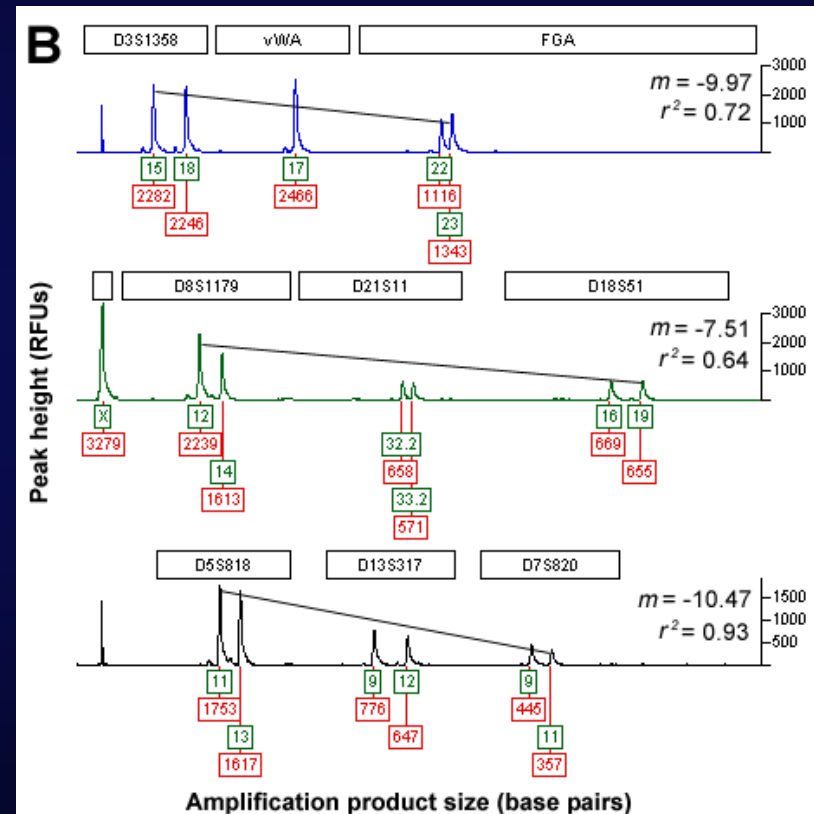


- Undegraded samples can have “ski-slopes” too.
- How negative does a slope have to be to an indication of degradation?
- Experience, training and expertise.
- Positive controls should not be degraded.

# Degradation, inhibition

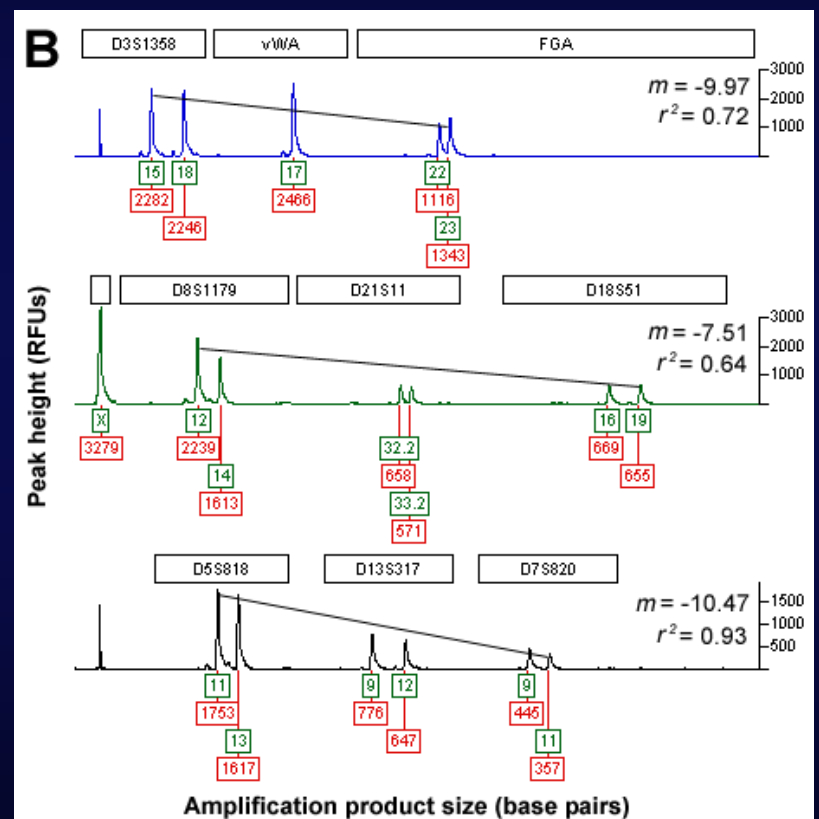
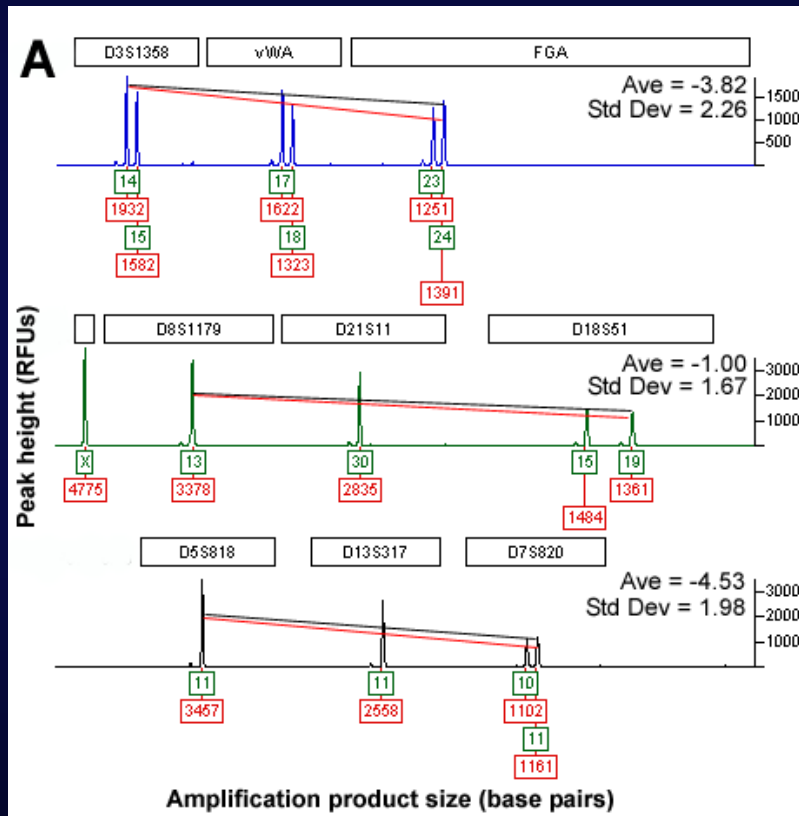
## The Leskie Inquest

- DNA profiles in a rape and a murder investigation match.
- Everyone agrees that the murder samples are degraded.
- If the rape sample is degraded, it could have contaminated the murder samples.
- Is the rape sample degraded?



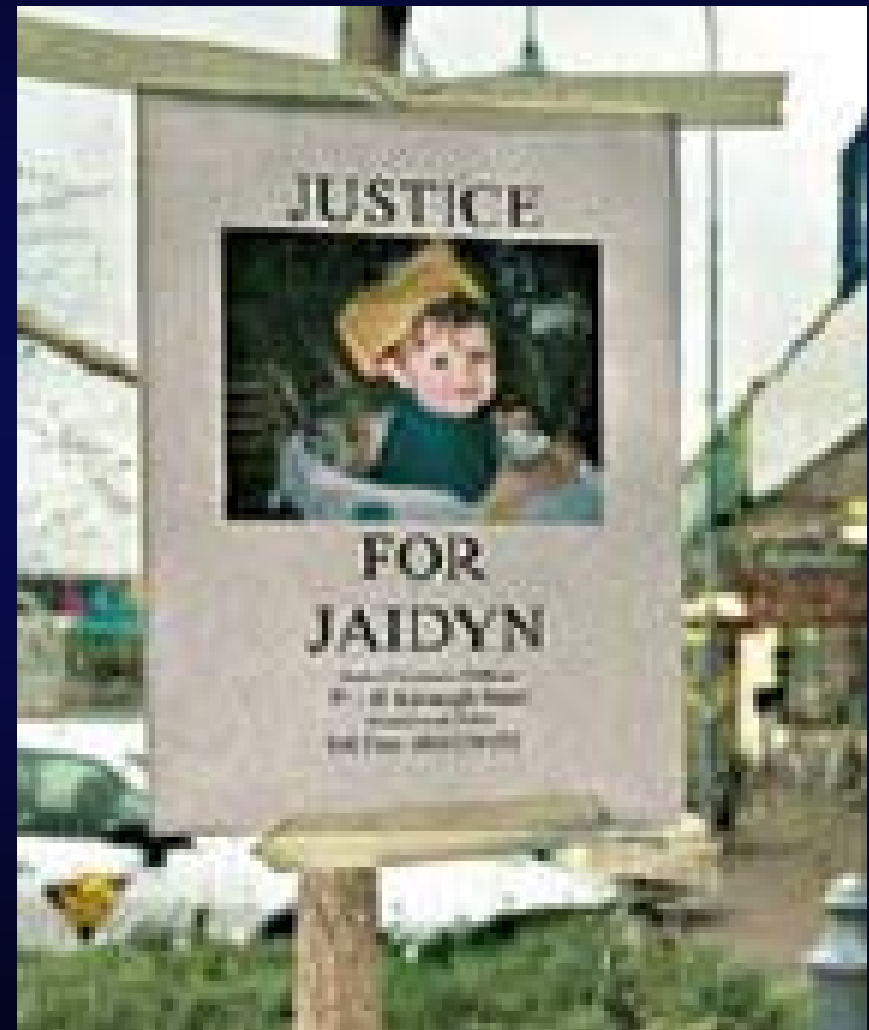
# Degradation, inhibition

## The Leskie Inquest



# Victorian Coroner's inquest into the death of Jaidyn Leskie

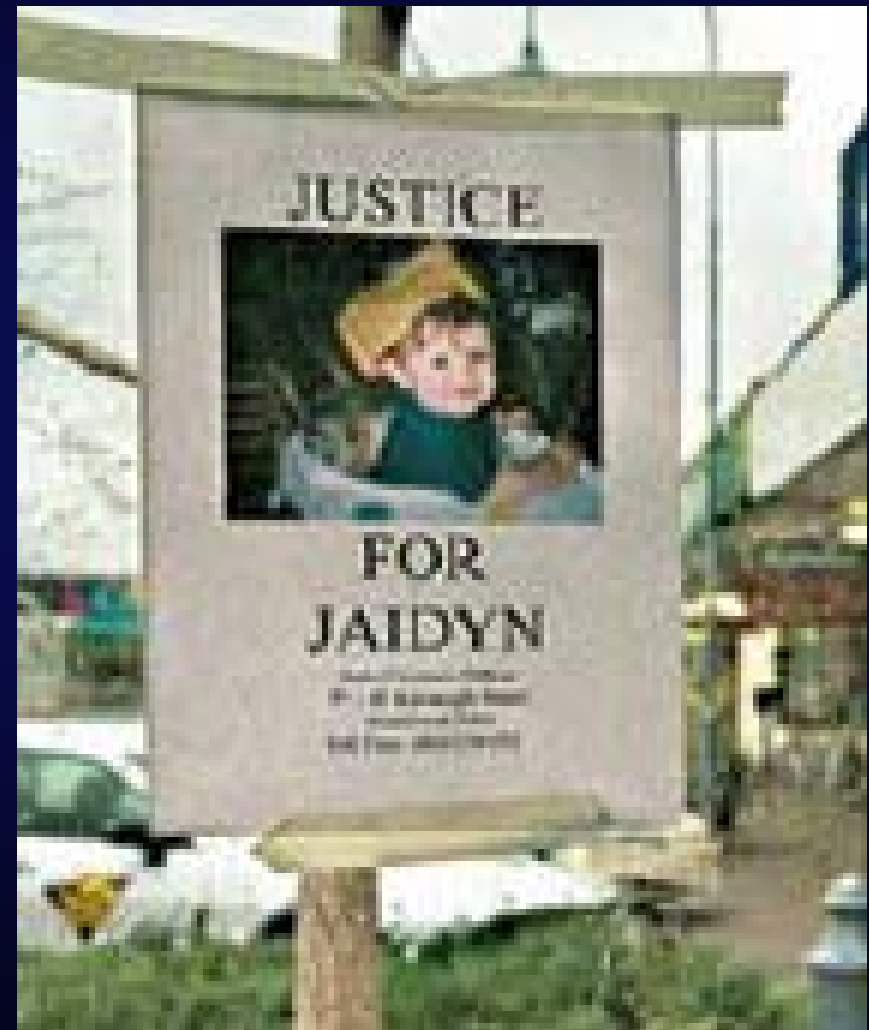
"8. During the conduct of the preliminary investigation (before it was decided to undertake an inquest) the female DNA allegedly taken from the bib that was discovered with the body was matched with a DNA profile in the Victorian Police Forensic Science database. This profile was from a rape victim who was subsequently found to be unrelated to the Leskie case."



# Victorian Coroner's inquest into the death of Jaidyn Leskie

"8. The match to the bib occurred as a result of contamination in the laboratory and was not an adventitious match. The samples from the two cases were examined by the same scientist within a close time frame."

[www.bioforensics.com/articles/Leskie\\_decision.pdf](http://www.bioforensics.com/articles/Leskie_decision.pdf)



The science of DNA profiling is  
sound.

But, not all of DNA profiling is  
science.

This is especially true in situations  
involving: small amounts of starting  
material, mixtures, relatives, and  
analyst judgment calls.



# Resources

- Internet
  - **Forensic Bioinformatics Website:** <http://www.bioforensics.com/>
  - **Applied Biosystems Website:** <http://www.appliedbiosystems.com/> (see human identity and forensics)
  - **STR base:** <http://www.cstl.nist.gov/biotech/strbase/> (very useful)
- Books
  - **'Forensic DNA Typing'** by John M. Butler (Academic Press)
- Scientists
  - **Larry Mueller** (UC Irvine)
  - **Simon Ford** (Lexigen, Inc. San Francisco, CA)
  - **William Shields** (SUNY, Syracuse, NY)
  - **Mike Raymer and Travis Doom** (Wright State, Dayton, OH)
  - **Marc Taylor** (Technical Associates, Ventura, CA)
  - **Keith Inman** (Forensic Analytical, Haywood, CA)
- Testing laboratories
  - **Technical Associates** (Ventura, CA)
  - **Forensic Analytical** (Haywood, CA)
- Other resources
  - **Forensic Bioinformatics** (Dayton, OH)